BUNDESREPUBLIK DEUTSCHLAND

PCT/EP04/ 11577
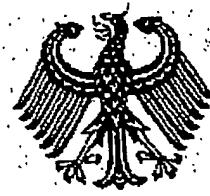
Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen:	103 48 407.8
Anmeldetag:	17. Oktober 2003
Anmelder/Inhaber:	Prof. Martin W i d s c h w e n d t e r , Mieming/AT
	Erstanmelder: Epigenomics AG, 10435 Berlin/DE
Bezeichnung:	Prognostic and Diagnostic Markers Markers for cell proliferative disorders of the breast tissues
IPC:	E 03 B 11/16

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 29. Dezember 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Kahle

Prognostic and Diagnostic Markers for cell proliferative disorders of the breast tissues

The present invention relates to prognostic and diagnostic markers for cell proliferative disorders of the breast tissues. The present invention therefore provides methods and nucleic acids for the analysis of biological samples for features associated with the development of breast cell proliferative disorders.

Background of the invention

This patent application relates to diagnosis and prognosis of breast cancer. Today involvement of axillary lymph nodes and tumour size are the most important prognostic factors in breast cancer. Although the presence or absence of metastatic involvement in the axillary lymph nodes is the most powerful prognostic factor available for patients with primary breast cancer, it is only an indirect measure reflecting the tumours' tendency to spread. In approximately one-third of women with breast cancer and negative lymph nodes the disease recurs, while about one-third of patients with positive lymph nodes are free of recurrence ten years after loco-regional therapy. These data highlight the need for more sensitive and specific prognostic indicators, ideally reflecting the presence or absence of tumour-specific alterations in the bloodstream that may eventually even after years lead to metastasis. It is now widely accepted that adjuvant systemic therapy substantially improves disease-free and overall survival in both pre- and postmenopausal women up to the age of 70 years with lymph node-negative or lymph node-positive breast cancer (Early Breast Cancer Trialists' Collaborative Group Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet, *351*: 1451-1467, 1998.2, 3). It is also generally accepted that patients with poor prognostic features benefit the most from adjuvant therapy, whereas some patients with good prognostic features may be overtreated (Goldhirsch et al.: Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. J. Clin. Oncol., *19*: 3817-3827, 2001.). Moreover many other factors have been investigated for their potential to predict disease outcome, but in general they have only limited predictive value. Recently, interesting prognostic parameters including gene-expression profiles, cell cycle regulating

proteins and occult cytokeratin-positive metastatic cells in the bone marrow have been added to the list of prognostic factors, but their prognostic relevance needs to be further evaluated.

Changes in the status of DNA methylation, known as epigenetic alterations, are one of the most common molecular alterations in human neoplasia, including breast cancer (Widschwendter and Jones: DNA methylation and breast carcinogenesis. Oncogene, *21*: 5462-5482, 2002). Cytosine methylation occurs after DNA synthesis by enzymatic transfer of a methyl group from the methyl donor S-adenosylmethionine to the carbon-5 position of cytosine. Cytosines are methylated in the human genome mostly when located 5′ to a guanosine. Regions with a high G:C content are so-called CpG islands. It has been increasingly recognized over the past four to five years that the CpG islands of a large number of genes, which are mostly unmethylated in normal tissue, are methylated to varying degrees in human cancers, thus representing tumor-specific alterations. The presence of abnormally high DNA concentrations in the serum of patients with various malignant diseases was described several years ago. The discovery that cell-free DNA can be shed into the bloodstream has generated great interest. Numerous studies have demonstrated tumor-specific alterations in DNA recovered from plasma or serum of patients with various malignancies, a finding that has potential for molecular diagnosis and prognosis. The nucleic acid markers described in plasma and serum include oncogene mutations, microsatellite alterations, gene rearrangements and epigenetic alterations, such as aberrant promoter hypermethylation (Anker et al.: Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. Cancer Metastasis Rev., *18*: 65-73, 1999). During recent years some studies have reported cell-free DNA in serum/plasma of breast cancer patients at diagnosis (for example: Silva et al.: Presence of tumor DNA in plasma of breast cancer patients: clinicopathological correlations. Cancer Res., *59*: 3251-3256, 1999) and in some cases persistence after primary therapy (for example: Silva et al.: Persistence of tumor DNA in plasma of breast cancer patients after mastectomy. Ann. Surg. Oncol., *9*: 71-76, 2002). Nevertheless an increasing number of studies have reported the presence of methylated DNA in serum/plasma of patients with various types of malignancies, including breast cancer, and the absence of methylated DNA in normal control patients (for example: Wong et al.: Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients. Cancer Res., *59*: 71-73, 1999). So far, only few studies have addressed the prognostic value of these epigenetic alterations in patients' bloodstream (Kawakami et al.: Hypermethylated APC DNA in plasma and prognosis of patients with esophageal adenocarcinoma. J. Natl. Cancer Inst., *92*: 1805-1811, 2000; Lecomte et al.: Detection of free-circulating

tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis. Int. J. Cancer, *100*: 542-548, 2002).

It will be appreciated by those skilled in the art that there exists a continuing need to improve methods of early detection, classification and treatment of breast cancers. In this application prognostic and diagnostic DNA methylation-based markers for breast cancer are disclosed.

5-methylcytosine positions cannot be identified by sequencing since 5-methylcytosine has the same base pairing behavior as cytosine. Moreover, the epigenetic information carried by 5-methylcytosine is completely lost during PCR amplification. Currently the most frequently used method for analysing DNA for 5-methylcytosine is based upon the specific reaction of bisulfite with cytosine which, upon subsequent alkaline hydrolysis, is converted to uracil which corresponds to thymidine in its base pairing behaviour. However, 5-methylcytosine remains unmodified under these conditions. Consequently, the original DNA is converted in such a manner that methylcytosine, which originally could not be distinguished from cytosine by its hybridisation behaviour, can now be detected as the only remaining cytosine using "normal" molecular biological techniques, for example, by amplification and hybridisation or sequencing. All of these techniques are based on base pairing which can now be fully exploited. In terms of sensitivity, the prior art is defined by a method which encloses the DNA to be analysed in an agarose matrix, thus preventing the diffusion and renaturation of the DNA (bisulfite only reacts with single-stranded DNA), and which replaces all precipitation and purification steps with fast dialysis (Olek A, Oswald J, Walter J. A modified and improved method for bisulphite based cytosine methylation analysis. Nucleic Acids Res. 1996 Dec 15;24(24):5064-6). Using this method, it is possible to analyse individual cells, which illustrates the potential of the method. However, currently only individual regions of a length of up to approximately 3000 base pairs are analysed, a global analysis of cells for thousands of possible methylation events is not possible. However, this method cannot reliably analyse very small fragments from small sample quantities either. These are lost through the matrix in spite of the diffusion protection.

An overview of the further known methods of detecting 5-methylcytosine may be gathered from the following review article: Fraga and Esteller: DNA Methylation: A Profile of Methods and Applications. Biotechniques 33:632-649, Sept. 2002.

To date, barring few exceptions (e.g., Zeschnigk M, Lich C, Buiting K, Doerfler W, Horsthemke B. A single-tube PCR test for the diagnosis of Angelman and Prader-Willi syndrome based on allelic methylation differences at the SNRPN locus. Eur J Hum Genet. 1997 Mar-Apr;5(2):94-8) the bisulfite technique is only used in research. Always, however, short, specific fragments of a known gene are amplified subsequent to a bisulfite treatment and either completely sequenced (Olek A, Walter J. The pre-implantation ontogeny of the H19 methylation imprint. Nat Genet. 1997 Nov;17(3):275-6) or individual cytosine positions are detected by a primer extension reaction (Gonzalgo ML, Jones PA. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). Nucleic Acids Res. 1997 Jun 15;25(12):2529-31, WO 95/00669) or by enzymatic digestion (Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res. 1997 Jun 15;25(12):2532-4). In addition, detection by hybridisation has also been described (Olek et al., WO 99/28498).

Further publications dealing with the use of the bisulfite technique for methylation detection in individual genes are: Grigg G, Clark S. Sequencing 5-methylcytosine residues in genomic DNA. Bioessays. 1994 Jun;16(6):431-6, 431; Zeschnigk M, Schmitz B, Dittrich B, Buiting K, Horsthemke B, Doerfler W. Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. Hum Mol Genet. 1997 Mar;6(3):387-95; Feil R, Charlton J, Bird AP, Walter J, Reik W. Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. Nucleic Acids Res. 1994 Feb 25;22(4):695-6; Martin V, Ribieras S, Song-Wang X, Rio MC, Dante R. Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. Gene. 1995 May 19;157(1-2):261-4; WO 97/46705, WO 95/15373, and WO 97/45560.

An overview of the Prior Art in oligomer array manufacturing can be gathered from a special edition of Nature Genetics (Nature Genetics Supplement, Volume 21, January 1999), published in January 1999, and from the literature cited therein.

Fluorescently labelled probes are often used for the scanning of immobilised DNA arrays. The simple attachment of Cy3 and Cy5 dyes to the 5'-OH of the specific probe are particularly suitable for fluorescence labels. The detection of the fluorescence of the hybridised probes

may be carried out, for example via a confocal microscope. Cy3 and Cy5 dyes, besides many others, are commercially available.

Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-TOF) is a very efficient development for the analysis of biomolecules (Karas M, Hillenkamp F. Laser desorption ionisation of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988 Oct 15;60(20):2299-301). An analyte is embedded in a light-absorbing matrix. The matrix is evaporated by a short laser pulse thus transporting the analyte molecule into the vapour phase in an unfragmented manner. The analyte is ionised by collisions with matrix molecules. An applied voltage accelerates the ions into a field-free flight tube. Due to their different masses, the ions are accelerated at different rates. Smaller ions reach the detector sooner than bigger ones.

MALDI-TOF spectrometry is excellently suited to the analysis of peptides and proteins. The analysis of nucleic acids is somewhat more difficult (Gut I G, Beck S. DNA and Matrix Assisted Laser Desorption Ionisation Mass Spectrometry. Current Innovations and Future Trends. 1995, 1; 147-57). The sensitivity to nucleic acids is approximately 100 times worse than to peptides and decreases disproportionally with increasing fragment size. For nucleic acids having a multiply negatively charged backbone, the ionisation process via the matrix is considerably less efficient. In MALDI-TOF spectrometry, the selection of the matrix plays an eminently important role. For the desorption of peptides, several very efficient matrixes have been found which produce a very fine crystallisation. There are now several responsive matrixes for DNA, however, the difference in sensitivity has not been reduced. The difference in sensitivity can be reduced by chemically modifying the DNA in such a manner that it becomes more similar to a peptide. Phosphorothioate nucleic acids in which the usual phosphates of the backbone are substituted with thiophosphates can be converted into a charge-neutral DNA using simple alkylation chemistry (Gut IG, Beck S. A procedure for selective DNA alkylation and detection by mass spectrometry. Nucleic Acids Res. 1995 Apr 25;23(8):1367-73). The coupling of a charge tag to this modified DNA results in an increase in sensitivity to the same level as that found for peptides. A further advantage of charge tagging is the increased stability of the analysis against impurities which make the detection of unmodified substrates considerably more difficult.

Genomic DNA is obtained from DNA of cell, tissue or other test samples using standard methods. This standard methodology is found in references such as Fritsch and Maniatis eds., Molecular Cloning: A Laboratory Manual, 1989.

Description

The present invention provides methods and nucleic acids for the analysis of biological samples for features associated with the development of breast cell proliferative disorders. The invention is characterised in that the nucleic acid of at least one member of the group of genes according to Table 1 is/are contacted with a reagent or series of reagents capable of distinguishing between methylated and non methylated CpG dinucleotides within the genomic sequence of interest. The present invention makes available a method for ascertaining genetic and/or epigenetic parameters of genomic DNA. The method is for use for the determining the prognosis of breast cell proliferative disorders. The invention presents improvements over the state of the art in that by means of the methods and compounds described herein a person skilled in the art may carry out a sensitive and specific detection assay of cellular matter comprising cancerous breast tissue. This is particularly useful as it allows the analysis of samples of body fluids which may contain only a minimal amount of cell proliferative disorder cellular matter, and enables the detection of said cells and the identification of the organ from which they originated (in this case breast). To date there are no known clinically utilisable means for the detection of breast cancer using genetic methylation markers to analyse bodily fluid samples, such as blood, lymphatic fluids, nipple aspirate and plasma. The generated information is useful in the selection of a treatment of the patient. If a positive prognosis is determined a further treatment might be redundant, while in a case of a poor prognosis a stronger treatment might be necessary.

Furthermore, the method enables the analysis of cytosine methylations and single nucleotide polymorphisms.

The genes that form the basis of the present invention are preferably to be used to form a "gene panel", i.e. a collection comprising the particular genetic sequences of the present invention and/or their respective informative methylation sites. The formation of gene panels allows for a quick and specific analysis of specific aspects of breast cancer. The gene panel(s) as described and employed in this invention can be used with surprisingly high efficiency for

the diagnosis, treatment and monitoring of and the analysis of a predisposition to breast cell proliferative disorders.

In addition, the use of multiple CpG sites from a diverse array of genes allows for a relatively high degree of sensitivity and specificity in comparison to single gene diagnostic and detection tools. Of the genes known to be specifically methylated in breast cancer, the particular combination of the genes according to the invention provides for a particularly sensitive and specific means for the identification of cell proliferative disorders of breast tissues.

The object of the invention is most preferably achieved by means of the analysis of the methylation patterns of one or a combination of genes taken from the group taken from the group ESR1, APC, HSD174B4, HIC1 and RASSF1A (see, for example, Table 1) and/or their regulatory regions. The invention is characterised in that the nucleic acid of one or a combination of genes taken from the group ESR1, APC, HSD174B4, HIC1 and RASSF1A are contacted with a reagent or series of reagents capable of distinguishing between methylated and non methylated CpG dinucleotides within the genomic sequence of interest.

The object of the invention can also be achieved by the analysis of the CpG methylation of one or a plurality of any subset of the group of genes ESR1, APC, HSD174B4, HIC1 and RASSF1A, in particular the following subsets are preferred:

.. RASSF1A and APC,
- RASSF1A, and
- APC

The present invention makes available a method for ascertaining genetic and/or epigenetic parameters of genomic DNA. The method is for use in the improved diagnosis, treatment and monitoring of breast cell proliferative disorders.

The disclosed invention further provides a method for determining the phenotype of a subject with a breast cell proliferative disorder comprising
a) obtaining a biological sample containing genomic DNA from said subject,
b) analysing the methylation pattern of one or more target nucleic acids comprising one or a combination of the genes taken from the group consisting of ESR1, APC, HSD174B4, HIC1

and RASSF1A and/or their regulatory regions by contacting at least one of said target nucleic acids in the biological sample obtained from said subject with at least one reagent, or series of reagents that distinguishes between methylated and non-methylated CpG dinucleotides, and

c) determining the phenotype of the individual by comparison to two known phenotypes, a first phenotype characterised by hypermethylation of the target nucleic acid and poor prognosis as relative to a second phenotype characterised by hypomethylation of the analysed target nucleic acid and better prognosis

The DNA may be obtained from any form of biological sample including but not limited to cell lines, histological slides, biopsies, tissue embedded in paraffin, breast tissues, blood, plasma, lymphatic fluid, lymphatic tissue, duct cells, ductal lavage fluid, nipple aspiration fluid and combinations thereof. Genomic DNA must then be isolated from the sample using any means standard in the art. The isolated DNA is treated with at least one reagent, or series of reagents that distinguishes between methylated and non-methylated CpG dinucleotides. This may be carried out by any means standard in the art including the use of restriction endonucleases. However, it is preferably carried out with bisulfite (sulfite, disulfite) and subsequent alkaline hydrolysis which results in a conversion of non-methylated cytosine nucleobases to uracil or to another base which is dissimilar to cytosine in terms of base pairing behaviour. If bisulfite solution is used for the reaction, then an addition takes place at the non-methylated cytosine bases. Moreover, a denaturating reagent or solvent as well as a radical interceptor must be present. A subsequent alkaline hydrolysis then gives rise to the conversion of non-methylated cytosine nucleobases to uracil. The converted DNA is then used for the detection of methylated cytosines. The methylation status of one or more of the genes ESR1, APC, HSD174B4, HIC1 and RASSF1Aand/or their regulatory regions is then analysed. This analysis may be carried out by any means standard in the art including the above described techniques. In the final step of the method the methylation pattern of the DNA obtained from the subject is compared to that of two known phenotypes. The first phenotype is characterised by hypermethylation or methylation of the target nucleic acid and poor prognosis as relative to a second phenotype characterised by hypomethylation or no methylation of the analysed target nucleic acid and better prognosis. It is particularly preferred that the genes APC and/RASSF1A are analysed. By determining which of the two phenotypes the subject belongs to it is possible to determine a suitable treatment to her breast cell proliferative disorder.

The method according to the invention may be used for the analysis of a wide variety of cell proliferative disorders of the breast tissues including, but not limited to, ductal carcinoma *in situ*, lobular carcinoma, colloid carcinoma, tubular carcinoma, medullary carcinoma, metaplastic carcinoma, intraductal carcinoma *in situ*, lobular carcinoma *in situ* and papillary carcinoma *in situ*.

Furthermore, the method enables the analysis of cytosine methylations and single nucleotide polymorphisms within said genes.

The object of the invention is achieved by means of the analysis of the methylation patterns of one or more of the genes ESR1, APC, HSD174B4, HIC1 and RASSF1A and/or their regulatory regions. In a particularly preferred embodiment the sequences of said genes comprise SEQ ID NOs: 1 to 5 and sequences complementary thereto.

The object of the invention may also be achieved by analysing the methylation patterns of one or more genes taken from the following subsets of said aforementioned group of genes. In one embodiment the object of the invention is achieved by analysis of the methylation patterns of the genes RASSF1A and APC and wherein it is further preferred that the sequence of said genes comprise SEQ ID NOs: 5 and 3, respectively. In a further embodiment the object of the invention is achieved by analysis of the methylation patterns of the gene RASSF1A and/or its regulatory sequences, and wherein it is further preferred that the sequence of said gene comprises SEQ ID NO: 5. The object of the invention may also be achieved by analysis of the methylation pattern of the gene APC and/or its regulatory sequences, and wherein it is further preferred that the sequence of said gene comprises SEQ ID NO: 3.

In a preferred embodiment said method is achieved by contacting said nucleic acid sequences in a biological sample obtained from a subject with at least one reagent or a series of reagents, wherein said reagent or series of reagents, distinguishes between methylated and non methylated CpG dinucleotides within the target nucleic acid.

In a preferred embodiment, the method comprises the following steps:
In the first step of the method the genomic DNA sample must be isolated from sources such as cells or cellular components which contain DNA, sources of DNA comprising, for example, cell lines, histological slides, biopsies, tissue embedded in paraffin, breast tissues, blood,

plasma, lymphatic fluid, lymphatic tissue, duct cells, ductal lavage fluid, nipple aspiration fluid and combinations thereof. Extraction may be by means that are standard to one skilled in the art, these include the use of detergent lysates, sonification and vortexing with glass beads. Once the nucleic acids have been extracted the genomic double stranded DNA is used in the analysis.

In one embodiment the DNA may be cleaved prior to the next step of the method, this may be by any means standard in the state of the art, in particular, but not limited to, with restriction endonucleases.

In the second step of the method, the genomic DNA sample is treated in such a manner that cytosine bases which are unmethylated at the 5'-position are converted to uracil, thymine, or another base which is dissimilar to cytosine in terms of hybridisation behaviour. This will be understood as "pretreatment" or "chemical pretreatment" hereinafter.

The above described treatment of genomic DNA is preferably carried out with bisulfite (sulfite, disulfite) and subsequent alkaline hydrolysis which results in a conversion of non-methylated cytosine nucleobases to uracil or to another base which is dissimilar to cytosine in terms of base pairing behaviour. If bisulfite solution is used for the reaction, then an addition takes place at the non-methylated cytosine bases. Moreover, a denaturating reagent or solvent as well as a radical interceptor must be present. A subsequent alkaline hydrolysis then gives rise to the conversion of non-methylated cytosine nucleobases to uracil. The converted DNA is then used for the detection of methylated cytosines.

Fragments of the pretreated DNA are amplified, using sets of primer oligonucleotides, and a preferably heat-stable, polymerase. Because of statistical and practical considerations, preferably more than six different fragments having a length of 100 - 2000 base pairs are amplified. The amplification of several DNA segments can be carried out simultaneously in one and the same reaction vessel. Usually, the amplification is carried out by means of a polymerase chain reaction (PCR).

The design of such primers is obvious to one skilled in the art. These should include at least two oligonucleotides whose sequences are each reverse complementary or identical to an at least 18 base-pair long segment of the following base sequences specified in the appendix:

SEQ ID NO 6 to 26. Said primer oligonucleotides are preferably characterised in that they do not contain any CpG dinucleotides. In a particularly preferred embodiment of the method, the sequence of said primer oligonucleotides are designed so as to selectively anneal to and amplify, only the breast cell specific DNA of interest, thereby minimising the amplification of background or non relevant DNA. In the context of the present invention, background DNA is taken to mean genomic DNA which does not have a relevant tissue specific methylation pattern, in this case, the relevant tissue being breast tissues.

According to the present invention, it is preferred that at least one primer oligonucleotide is bound to a solid phase during amplification. The different oligonucleotide and/or PNA-oligomer sequences can be arranged on a plane solid phase in the form of a rectangular or hexagonal lattice, the solid phase surface preferably being composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold, it being possible for other materials such as nitrocellulose or plastics to be used as well.

The fragments obtained by means of the amplification may carry a directly or indirectly detectable label. Preferred are labels in the form of fluorescence labels, radionuclides, or detachable molecule fragments having a typical mass which can be detected in a mass spectrometer, it being preferred that the fragments that are produced have a single positive or negative net charge for better detectability in the mass spectrometer. The detection may be carried out and visualised by means of matrix assisted laser desorption/ionisation mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI).

In the next step the nucleic acid amplificates are analysed in order to determine the methylation status of the genomic DNA prior to treatment.

The post treatment analysis of the nucleic acids may be carried out using alternative methods. Several methods for the methylation status specific analysis of the treated nucleic acids are described below, other alternative methods will be obvious to one skilled in the art.

The analysis may be carried out during the amplification step of the method. In one such embodiment, the methylation status of preselected CpG positions within the genes ESR1, APC, HSD174B4, HIC1 and RASSF1A and/or their regulatory regions may be detected by use of methylation specific primer oligonucleotides. The term "MSP" (Methylation-specific PCR)

refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996, and also disclosed in US Patents No. 5,786,146 and No. 6,265,171. The use of methylation status specific primers for the amplification of bisulphite treated DNA allows the differentiation between methylated and unmethylated nucleic acids. MSP primers pairs contain at least one primer which hybridises to a bisulphite treated CpG dinucleotide. Therefore the sequence of said primers comprises at least one CG, TG or CA dinucleotide. MSP primers specific for non methylated DNA contain a 'T' at the 3' position of the C position in the CpG. According to the present invention, it is therefore preferred that the base sequence of said primers is required to comprise a sequence having a length of at least 10 nucleotides which hybridises to a pretreated nucleic acid sequence according to SEQ ID NOs.: 6 to 26 and sequences complementary thereto wherein the base sequence of said oligomers comprises at least one CG, TG or CA dinucleotide.

In one embodiment of the method the methylation status of the CpG positions may be determined by means of hybridisation analysis. In this embodiment of the method the amplificates obtained in the second step of the method are hybridised to an array or a set of oligonucleotides and/or PNA probes. In this context, the hybridisation takes place in the manner described as follows. The set of probes used during the hybridisation is preferably composed of at least 4 oligonucleotides or PNA-oligomers. In the process, the amplificates serve as probes which hybridise to oligonucleotides previously bonded to a solid phase. The non-hybridised fragments are subsequently removed. Said oligonucleotides contain at least one base sequence having a length of 10 nucleotides which is reverse complementary or identical to a segment of the base sequences specified in the appendix, the segment containing at least one CpG or TpG dinucleotide. In a further preferred embodiment the cytosine of the CpG dinucleotide, or in the case of TpG, the thiamine, is the 5th to 9th nucleotide from the 5'-end of the 10-mer. One oligonucleotide exists for each CpG or TpG dinucleotide.

The non-hybridised amplificates are then removed. In the final step of the method, the hybridised amplificates are detected. In this context, it is preferred that labels attached to the amplificates are identifiable at each position of the solid phase at which an oligonucleotide sequence is located.

In a preferred embodiment of the method the methylation status of the CpG positions may be determined by means of oligonucleotide probes that are hybridised to the treated DNA con-

currently with the PCR amplification primers (wherein said primers may either be methylation specific or standard).

A particularly preferred embodiment of this method is the use of fluorescence-based Real Time Quantitative PCR (Heid et al., Genome Res. 6:986-994, 1996) employing a dual-labelled fluorescent oligonucleotide probe (TaqMan™ PCR, using an ABI Prism 7700 Sequence Detection System, Perkin Elmer Applied Biosystems, Foster City, California). The TaqMan™ PCR reaction employs the use of a nonextendible interrogating oligonucleotide, called a TaqMan™ probe, which is designed to hybridise to a GpC-rich sequence located between the forward and reverse amplification primers. The TaqMan™ probe further comprises a fluorescent "reporter moiety" and a "quencher moiety" covalently bound to linker moieties (e.g., phosphoramidites) attached to the nucleotides of the TaqMan™ oligonucleotide. For analysis of methylation within nucleic acids subsequent to bisulphite treatment it is required that the probe be methylation specific, as described in U.S. 6,331,393, also known as the Methyl Light assay. Variations on the TaqMan™ detection methodology that are also suitable for use with the described invention include the use of dual probe technology (Lightcycler™) or fluorescent amplification primers (Sunrise™ technology). Both these techniques may be adapted in a manner suitable for use with bisulphite treated DNA, and moreover for methylation analysis within CpG dinucleotides.

A further suitable method for the use of probe oligonucleotides for the assessment of methylation by analysis of bisulphite treated nucleic acids is the use of blocker oligonucleotides. The use of such oligonucleotides has been described by D. Yu, M. Mukai, Q. Liu, C. Steinman in BioTechniques 23(4), 1997, 714-720. Blocking probe oligonucleotides are hybridised to the bisulphite treated nucleic acid concurrently with the PCR primers. PCR amplification of the nucleic acid is terminated at the 5' position of the blocking probe, thereby amplification of a nucleic acid is suppressed wherein the complementary sequence to the blocking probe is present. The probes may be designed to hybridise to the bisulphite treated nucleic acid in a methylation status specific manner. For example, for detection of methylated nucleic acids within a population of unmethylated nucleic acids suppression of the amplification of nucleic acids which are unmethylated at the position in question would be carried out by the use of blocking probes comprising a 'CG' at the position in question, as opposed to a 'CA'.

For PCR methods using blocker oligonucleotides, efficient disruption of polymerase-mediated amplification requires that blocker oligonucleotides not be elongated by the polymerase. Preferably, this is achieved through the use of blockers that are 3'-deoxyoligonucleotides, or oligonucleotides derivatised at the 3' position with other than a "free" hydroxyl group. For example, 3'-O-acetyl oligonucleotides are representative of a preferred class of blocker molecule.

Additionally, polymerase-mediated decomposition of the blocker oligonucleotides should be precluded. Preferably, such preclusion comprises either use of a polymerase lacking 5'-3' exonuclease activity, or use of modified blocker oligonucleotides having, for example, thioate bridges at the 5'-terminii thereof that render the blocker molecule nuclease-resistant. Particular applications may not require such 5' modifications of the blocker. For example, if the blocker- and primer-binding sites overlap, thereby precluding binding of the primer (e.g., with excess blocker), degradation of the blocker oligonucleotide will be substantially precluded. This is because the polymerase will not extend the primer toward, and through (in the 5'-3' direction) the blocker - a process that normally results in degradation of the hybridized blocker oligonucleotide.

A particularly preferred blocker/PCR embodiment, for purposes of the present invention and as implemented herein, comprises the use of peptide nucleic acid (PNA) oligomers as blocking oligonucleotides. Such PNA blocker oligomers are ideally suited, because they are neither decomposed nor extended by the polymerase.

Preferably, therefore, the base sequence of said blocking oligonucleotides is required to comprise a sequence having a length of at least 9 nucleotides which hybridises to a pretreated nucleic acid sequence according to one of SEQ ID NOs: 6 to 26 and sequences complementary thereto, wherein the base sequence of said oligonucleotides comprises at least one CpG, TpG or CpA dinucleotide.

In a further preferred embodiment of the method the determination of the methylation status of the CpG positions is carried out by the use of template directed oligonucleotide extension, such as MS SNuPE as described by Gonzalgo and Jones (Nucleic Acids Res. 25:2529-2531).

In a further embodiment of the method the determination of the methylation status of the CpG positions is enabled by sequencing and subsequent sequence analysis of the amplificate generated in the second step of the method (Sanger F., et al., 1977 PNAS USA 74: 5463-5467).

The method according to the invention may be enabled by any combination of the above means. In a particularly preferred mode of the invention the use of real time detection probes is concurrently combined with MSP and/or blocker oligonucleotides.

A further embodiment of the invention is a method for the analysis of the methylation status of genomic DNA without the need for pretreatment. In the first and second steps of the method the genomic DNA sample must be obtained and isolated from tissue or cellular sources. Such sources may include cell lines, histological slides, biopsies, tissue embedded in paraffin, breast tissues, blood, plasma, lymphatic fluid, lymphatic tissue, duct cells, ductal lavage fluid, nipple aspiration fluid and combinations thereof. Extraction may be by means that are standard to one skilled in the art, these include the use of detergent lysates, sonication and vortexing with glass beads. Once the nucleic acids have been extracted the genomic double stranded DNA is used in the analysis.

In a preferred embodiment the DNA may be cleaved prior to the treatment, this may be by any means standard in the state of the art, in particular with restriction endonucleases. In the third step, the DNA is then digested with one or more methylation sensitive restriction enzymes. The digestion is carried out such that hydrolysis of the DNA at the restriction site is informative of the methylation status of a specific CpG dinucleotide.

In a preferred embodiment the restriction fragments are amplified. In a further preferred embodiment this is carried out using the polymerase chain reaction.

In the final step the amplificates are detected. The detection may be by any means standard in the art, for example, but not limited to, gel electrophoresis analysis, hybridisation analysis, incorporation of detectable tags within the PCR products, DNA array analysis, MALDI or ESI analysis.

The aforementioned method is preferably used for ascertaining genetic and/or epigenetic parameters of genomic DNA.

In order to enable this method, the invention further provides the modified DNA of one or a combination of genes taken from the group ESR1, APC, HSD174B4, HIC1 and RASSF1A as well as oligonucleotides and/or PNA-oligomers for detecting cytosine methylations within said genes. The present invention is based on the discovery that genetic and epigenetic parameters and, in particular, the cytosine methylation patterns of said genomic DNAs are particularly suitable for improved treatment and monitoring of breast cell proliferative disorders.

The nucleic acids according to the present invention can be used for the analysis of genetic and/or epigenetic parameters of genomic DNA.

In another aspect of the present invention, the object of the present invention is achieved using a nucleic acid containing a sequence of at least 18 bases in length of the pretreated genomic DNA according to one of SEQ ID NOs: 6 to 25 and sequences complementary thereto.

The modified nucleic acids could heretofore not be connected with the ascertainment of disease relevant genetic and epigenetic parameters.

The object of the present invention is further achieved by an oligonucleotide or oligomer for the analysis of pretreated DNA, for detecting the genomic cytosine methylation state, said oligonucleotide containing at least one base sequence having a length of at least 10 nucleotides which hybridises to a pretreated genomic DNA according to SEQ ID Nos: 6 to 26. The oligomer probes according to the present invention constitute important and effective tools which, for the first time, make it possible to ascertain specific genetic and epigenetic parameters during the analysis of biological samples for features associated with a patient's response to endocrine treatment. Said oligonucleotides allow the improved treatment and monitoring of breast cell proliferative disorders. The base sequence of the oligomers preferably contains at least one CpG or TpG dinucleotide. The probes may also exist in the form of a PNA (peptide nucleic acid) which has particularly preferred pairing properties. Particularly preferred are oligonucleotides according to the present invention in which the cytosine of the CpG dinucleotide is within the middle third of said oligonucleotide e.g. the 5th - 9th nucleotide from the 5'-end of a 13-mer oligonucleotide; or in the case of PNA-oligomers, it is preferred for the cytosine of the CpG dinucleotide to be the 4th - 6th nucleotide from the 5'-end of the 9-mer.

The oligomers according to the present invention are normally used in so called "sets" which contain at least two oligomers and up to one oligomer for each of the CpG dinucleotides within SEQ ID NOs: 6 to 26.

In the case of the sets of oligonucleotides according to the present invention, it is preferred that at least one oligonucleotide is bound to a solid phase. It is further preferred that all the oligonucleotides of one set are bound to a solid phase.

The present invention further relates to a set of at least 2 n (oligonucleotides and/or PNA-oligomers) used for detecting the cytosine methylation state of genomic DNA, by analysis of said sequence or treated versions of said sequence (of the genes ESR1, APC, HSD174B4, HIC1 and RASSF1A, as detailed in the sequence listing and Table 1) and sequences comple-mentary thereto). These probes enable improved treatment and monitoring of breast cell pro-liferative disorders.

The set of oligomers may also be used for detecting single nucleotide polymorphisms (SNPs) by analysis of said sequence or treated versions of said sequence of the genes ESR1, APC, HSD174B4, HIC1 and RASSF1A .

It will be obvious to one skilled in the art that the method according to the invention will be improved and supplemented by the incorporation of markers and clinical indicators known in the state of the art and currently used as diagnostic or prognostic markers. More preferably said markers include node status, age, menopausal status, grade, estrogen and progesterone receptors.

The genes that form the basis of the present invention may be used to form a "gene panel", i.e. a collection comprising the particular genetic sequences of the present invention and/or their respective informative methylation sites. The formation of gene panels allows for a quick and specific analysis of specific aspects of breast cancer treatment. The gene panel(s) as described and employed in this invention can be used with surprisingly high efficiency for the treatment of breast cell proliferative disorders by prediction of the outcome of treatment with a therapy comprising one or more drugs which target the estrogen receptor pathway or are involved in estrogen metabolism, production, or secretion. The analysis of each gene of the panel contrib-utes to the evaluation of patient responsiveness, however, in a less prefered embodiment the

patient evaluation may be achieved by analysis of only a single gene. The analysis of a single member of the 'gene panel' would enable a cheap but less accurate means of evaluating patient responsiveness, the analysis of multiple members of the panel would provide a rather more expensive means of carrying out the method, but with a higher accuracy (the technically preferred solution).

According to the present invention, it is preferred that an arrangement of different oligonucleotides and/or PNA-oligomers (a so-called "array") made available by the present invention is present in a manner that it is likewise bound to a solid phase. This array of different oligonucleotide- and/or PNA-oligomer sequences can be characterised in that it is arranged on the solid phase in the form of a rectangular or hexagonal lattice. The solid phase surface is preferably composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold. However, nitrocellulose as well as plastics such as nylon which can exist in the form of pellets or also as resin matrices are suitable alternatives.

Therefore, a further subject matter of the present invention is a method for manufacturing an array fixed to a carrier material for the improved treatment and monitoring of breast cell proliferative disorders. In said method at least one oligomer according to the present invention is coupled to a solid phase. Methods for manufacturing such arrays are known, for example, from US Patent 5,744,305 by means of solid-phase chemistry and photolabile protecting groups.

A further subject matter of the present invention relates to a DNA chip for the improved treatment and monitoring of breast cell proliferative disorders. The DNA chip contains at least one nucleic acid according to the present invention. DNA chips are known, for example, in US Patent 5,837,832.

Moreover, a subject matter of the present invention is a kit which may be composed, for example, of a bisulfite-containing reagent, a set of primer oligonucleotides containing at least two oligonucleotides whose sequences in each case correspond to or are complementary to a 18 base long segment of the base sequences specified in SEQ ID NOs: 6 to 26 and/or PNA-oligomers as well as instructions for carrying out and evaluating the described method.

In a further preferred embodiment said kit may further comprise standard reagents for performing a CpG position specific methylation analysis wherein said analysis comprises one or more of the following techniques: MS-SNuPE, MSP, Methyl light, Heavy Methyl, and nucleic acid sequencing. However, a kit along the lines of the present invention can also contain only part of the aforementioned components.

Typical reagents (e.g., as might be found in a typical MethyLight®-based kit) for MethyLight® analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); TaqMan® probes; optimized PCR buffers and deoxynucleotides; and Taq polymerase.

Typical reagents (e.g., as might be found in a typical Ms-SNuPE-based kit) for Ms-SNuPE analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimized PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for specific gene; reaction buffer (for the Ms-SNuPE reaction); and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery regents or kit (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

Typical reagents (e.g., as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

The oligomers according to the present invention or arrays thereof as well as a kit according to the present invention are intended to be used for the improved treatment and monitoring of breast cell proliferative disorders. According to the present invention, the method is preferably used for the analysis of important genetic and/or epigenetic parameters within genomic DNA, in particular for use in improved treatment and monitoring of breast cell proliferative disorders.

The methods according to the present invention are used, for improved detection, treatment and monitoring of breast cell proliferative disorder.

The present invention moreover relates to the diagnosis and/or prognosis of events which are disadvantageous or relevant to patients or individuals in which important genetic and/or epigenetic parameters within genomic DNA, said parameters obtained by means of the present invention may be compared to another set of genetic and/or epigenetic parameters, the differences serving as the basis for the diagnosis and/or prognosis of events which are disadvantageous or relevant to patients or individuals.

In the context of the present invention the term "hybridisation" is to be understood as a bond of an oligonucleotide to a completely complementary sequence along the lines of the Watson-Crick base pairings in the sample DNA, forming a duplex structure.

In the context of the present invention, "genetic parameters" are mutations and polymorphisms of genomic DNA and sequences further required for their regulation. To be designated as mutations are, in particular, insertions, deletions, point mutations, inversions and polymorphisms and, particularly preferred, SNPs (single nucleotide polymorphisms).

In the context of the present invention the term "methylation state" is taken to mean the degree of methylation present in a nucleic acid of interest, this may be expressed in absolute or relative terms i.e. as a percentage or other numerical value or by comparison to another tissue and therein described as hypermethylated, hypomethylated or as having significantly similar or identical methylation status.

In the context of the present invention the term "regulatory region" of a gene is taken to mean nucleotide sequences which affect the expression of a gene. Said regulatory regions may be located within, proximal or distal to said gene. Said regulatory regions include but are not limited to constitutive promoters, tissue-specific promoters, developmental-specific promoters, inducible promoters and the like. Promoter regulatory elements may also include certain enhancer sequence elements that control transcriptional or translational efficiency of the gene.

In the context of the present invention the term "chemotherapy" is taken to mean the use of drugs or chemical substances to treat cancer. This definition excludes radiation therapy (treatment with high energy rays or particles), hormone therapy (treatment with hormones or hormone analogues (synthetic substitutes) and surgical treatment.

In the context of the present invention, "epigenetic parameters" are, in particular, cytosine methylations and further modifications of DNA bases of genomic DNA and sequences further required for their regulation. Further epigenetic parameters include, for example, the acetylation of histones which, cannot be directly analysed using the described method but which, in turn, correlates with the DNA methylation.

In the following, the present invention will be explained in greater detail on the basis of the sequences, figures and examples without being limited thereto.

Figure 1 shows the Kaplan-Meier estimated overall survival curves for the gene APC, for a set of 86 breast cancer patients. The dotted line (upper curve) shows unmethylated samples whereas the unbroken line (lower curve) shows methylated samples. The x-axis shows the number of years, and the Y-axis shows the proportion of the group.

Figure 2 shows the Kaplan-Meier estimated overall survival curves for the gene RASSF1A, for a set of 86 breast cancer patients. The dotted line (upper curve) shows unmethylated samples whereas the unbroken line (lower curve) shows methylated samples. The x-axis shows the number of years, and the Y-axis shows the proportion of the group.

Figure 3 shows the combined Kaplan-Meier estimated overall survival curves for the genes APC and/or RASSF1A, for a set of 86 breast cancer patients. The dotted line (upper curve) shows unmethylated samples whereas the unbroken line (lower curve) shows methylated samples. The x-axis shows the number of years, and the Y-axis shows the proportion of the group.

SEQ ID NOs: 1 to 5 represent 5' and/or regulatory regions and/or CpG rich regions of the genes according to Table 1. These sequences are derived from Genbank and will be taken to include all minor variations of the sequence material which are currently unforeseen, for example, but not limited to, minor deletions and SNPs.

SEQ ID NOs: 6 to 26 exhibit the pretreated sequence of DNA derived from the genomic sequence according to Table 1. These sequences will be taken to include all minor variations of

the sequence material which are currently unforeseen, for example, but not limited to, minor deletions and SNPs.

Examples

Using MethyLight, a high-throughput DNA methylation assay, the inventors analysed 39 genes in a gene evaluation set, consisting of ten sera from metastasised patients, 26 patients with primary breast cancer and ten control patients. In order to determine the prognostic value of genes identified within the gene evaluation set, the inventors finally analysed pretreatment sera of 24 patients having had no adjuvant treatment (training set) to determine their prognostic value. An independent test set consisting of 62 patients was then used to test the validity of genes and combinations of genes, which in the training set were found to be good prognostic markers.

In the gene evaluation set the inventors identified five genes (ESR1, APC, HSD17B4, HIC1 and RASSF1A). In the training set, patients with methylated serum DNA for RASSF1A and/or APC had the worst prognosis (p < 0.001). This finding was confirmed by analysing serum samples from the independent test set (p = 0.007). When analysing all 86 investigated patients, multivariate analysis showed methylated RASSF1A and/or APC serum DNA to be independently associated with poor outcome, with a relative risk for death of 5.7. DNA methylation of particular genes in pretherapeutic sera of breast cancer patients, especially of RASSF1A/APC, is more powerful than standard prognostic parameters.

The gene evaluation set consisted of patients with recurrent disease (n=10; sera obtained at diagnosis of metastasis in the bone, lung, brain or liver) and pretherapeutic sera of recently diagnosed primary breast cancer patients (n=26; age range: 36.1 yrs to 83.9 yrs. (mean: 59.3 yrs.); two, 18 and six patients had pT1, pT2 and pT3 cancers, respectively; 15, ten and one patients had lymph node-negative, - positive and unknown disease, respectively) and normal controls (n=10; age range: 20.5 to 71.5 yrs. (mean: 44.6 yrs.); all underwent a core biopsy and were confirmed to have benign disease of the breast).

To assess prognostic significance the inventors used pretherapeutic sera in independent training (n=24) and test (n=62) sets consisting of patients who did not receive any adjuvant treatment after surgery.

Systemic adjuvant therapy was either not necessary or the patients were not eligible or refused any further treatment. The primary surgical procedure included breast-conserving lumpec-tomy or modified radical mastectomy and axillary lymph node dissection. Median age of the study population was 60 years (range, 28 to 86 yrs.). After a median follow-up of 3.7 yrs. (range: one month to 12.2 yrs.) 17 of the 86 patients (20 %) had died. Distribution of aberrant serum DNA methylation of the 86 patients and association with clinical and histopathological characteristics are shown in Table 2.

Patients' blood samples were drawn prior to therapeutic intervention. The blood was centri-fuged at 2000 g for 10 min at room temperature, and 1 mL aliquots of serum samples were stored at –30°C. Genomic DNA from serum samples was isolated using the High Pure Viral Nucleic Acid Kit (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocol with some modifications for multiple loading of the DNA extraction columns to gain a sufficient amount of DNA. Thus, 4 x 200µl of a serum sample were each mixed with 200µl of working solution (binding buffer supplemented with polyA carrier RNA) and 50µl protei-nase K [18 mg/ml] and incubated for 10 minutes at 72°C. After adding 100µl isopropanol the solution was mixed, loaded onto the extraction column and centrifuged for 1 minute at 8000g. The flow-through was pipetted back into the same column reservoir and centrifuged a second time. This procedure was repeated four times for each serum sample. After these "pooling steps" the DNA isolation was processed as described in the manufacturer's protocol. For DNA elution 55µl of AE-buffer (Quiagen, CA, USA) were added, incubated for 20 min at 45°C and centrifuged for three minutes at 12.000g. For both, normal sera and cancer sera analysis the same amount of serum for DNA extraction was used.

Sodium bisulfite conversion of genomic DNA was performed as described previously (Eads et al.: MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res., 28: E32, 2000).

Sodium bisulfite-treated genomic DNA was analysed by means of the MethyLight, a fluores-cence-based, real-time PCR assay, as described previously (Eads et al 2000, see above, Eads et al.: Epigenetic patterns in the progression of esophageal adenocarcinoma. Cancer Res., 61: 3410-3418, 2001). Briefly, two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set for the gene of interest and a reference set, ß-

actin (*ACTB*), to normalise for input DNA. Serum samples of patients with recurrent disease revealed the highest amount of ß-actin, whereas no difference between ß-actin values from serum samples of patients with primary breast cancer and sera of normal controls was observed. Specificity of the reactions for methylated DNA was confirmed separately using *SssI* (New England Biolabs)-treated human white blood cell DNA (heavily methylated). The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE:ACTB* ratio of a sample by the *GENE:ACTB* ratio of *SssI*-treated white blood cell DNA and multiplying by 100. The abbreviation PMR (percentage of fully methylated reference) indicates this measurement. For each MethyLight reaction 10µl of bisulfite-treated genomic DNA was used.

A gene was deemed methylated if the PMR value was > 0. Primer and probes specific for methylated DNA and used for MethyLight reactions are listed in Supplemental Data.

The inventors used Pearsons Chi² or – in the case of low frequencies per cell - Fisher's exact method to test associations between categorically clinicopathological features. The Mann-Whitney-U-Test was used to assess differences between non-parametric distributed variables. Overall survival was calculated from the date of diagnosis of the primary tumour to the date of death or last follow-up. Overall survival curves were calculated with the Kaplan-Meier method. Univariate analysis of overall survival according to clinicopathological factors (histological type, tumour stage, nodal status, grading, menopausal status, hormone receptor status (estrogen and/or progesterone receptor positivity), estrogen and progesterone receptor status) and gene methylation were performed using a two-sided log-rank test.

Multivariate Cox proportional hazards analysis was used to estimate the prognostic effect of methylated genes.

A p value < 0.05 was considered a statistically significant difference. All statistical analyses were performed using SPSS Software 10.0.

The inventors initially investigated 39 genes in the sera of ten patients with metastasised breast cancer for the presence of aberrant methylation. The 33 genes positive in the sera of the metastasised patients were further evaluated in an independent sample set of pretherapeutic sera of 26 patients with primary breast cancer and ten healthy controls. An overview of the

frequency of methylation in the investigated serum samples is given in Table 3. The most appropriate genes for our further analyses were determined to be those that met one of the following criteria: (i) unmethylated in serum samples from healthy controls and > 10% methylated in serum samples from primary breast cancer patients, or (ii) ≤ 10% methylated in serum samples from healthy controls and >20% methylated in serum samples from primary breast cancer patients. A total of five genes, namely ESR1, APC, HSD17B4, HIC1 and RASSF1A, met at least one of these criteria (Table 3).

Pre-treatment serum samples from patients included in the training set were used to evaluate the prognostic value of the methylation status of these five genes. In this training set the inventors identified ESR1, APC or RASSF1A methylation in primary breast cancer patients' sera to be markers of poor prognosis, whereas HSD17B4 reached only borderline significance and aberrant methylation of HIC1 showed no significant results (Table 4). Furthermore, various combinations of the investigated genes were analyzed. Patients were classified as methylation-positive if at least one of the genes included in the combination showed aberrant methylation. Patients with methylated serum DNA for RASSF1A and/or APC had the worst prognosis (P < 0.001), even worse than when each gene was analysed individually (Table 4).

The highly significant prognostic value for APC and/or RASSF1A methylation in serum samples from breast cancer patients was confirmed by analysing the test set (P = 0.007, log rank test). ESR1 and APC methylation as single genes or the combinations ESR1/RASSF1A and ESR1/APC no longer had prognostic significance (Table 4).

Combined analysis of the training and test sets (n=86) showed correlation between ESR1 and RASSF1A (P = 0.005) and between ESR1 and APC (P = 0.031), whereas no correlation was observed between RASSF1A and APC. In patients with advanced tumours RASSF1A and ESR1 methylation and in patients with progesterone receptor-negative tumours APC methylation was more prevalent in pretherapeutic sera, while no further associations were seen between clinicopathological features and DNA methylation of APC, ESR1 or RASSF1A (Table 5). RASSF1A methylation in pretherapeutic sera was more prevalent in older than in younger patients, whereas age had no effect on DNA methylation of ESR1 or APC.

Univariate analysis of all 86 investigated patients (training set plus test set) revealed prognostic significance for tumour size, lymph node metastases and methylation status of APC,

RASSF1A and the combination of RASSF1A/APC (Table 6; Fig. 1). Due to the fact that ESR1 methylation correlates with APC as well as with RASSF1A methylation, the inventors did not test the triple combination in the univariate or the multivariate analyses of all 86 patients.

The Cox multiple-regression analysis included tumour size, lymph node metastases, age and methylation status of the investigated genes. Beside lymph node status, methylated RASSF1A and/or APC serum DNA was strongly associated with poor outcome, with a relative risk for death of 5.7 (Table 7).

Prognosis in patients with newly diagnosed breast cancer is determined primarily by the presence or absence of metastases in draining axillary lymph nodes. Nevertheless, the life-threatening event in breast cancer is not lymph node metastasis per se, but haematogenous metastases which mainly affect bone, liver, lung and brain. The inventors therefore aimed to develop a prognostic test that is sensitive for haematogenous metastases and could be performed in patients' pretherapeutic serum.

In recent years several studies have reported cell-free tumour-specific DNA in serum/plasma of breast cancer patients at diagnosis. Aberrant methylation of serum/plasma DNA of patients with various types of malignancies, including breast cancer, has been described (see above).

In light of these observations, the inventors examined the methylation status of 39 genes, which, on the one hand, are known to be frequently methylated in breast cancer and other malignancies (Jones and Baylin: The fundamental role of epigenetic events in cancer. Nat. Rev. Genet., 3: 415-428, 2002; Widschwendter and Jones: DNA methylation and breast carcinogenesis. Oncogene, 21: 5462-5482, 2002) and, on the other hand, were reported to be abnormally regulated in tumours of patients with poor prognostic breast cancer (van't Veer et al.: Gene expression profiling predicts clinical outcome of breast cancer. Nature, 415: 530-536, 2002; van de Vijver et al.: A gene-expression signature as a predictor of survival in breast cancer. N. Engl. J. Med., 347: 1999-2009, 2002) Because levels of circulating DNA in metastasised patients are known to be higher (Leon et al.: Free DNA in the serum of cancer patients and the effect of therapy Cancer Res., 37: 646-650, 1977) and because the loss of genetic heterogeneity of disseminated tumour cells with the emergence of clinically evident metastasis was recently reported (Klein et al.: Genetic heterogeneity of single disseminated

tumour cells in minimal residual cancer. Lancet, 360: 683-689, 2002), the inventors firstly investigated these 39 genes in ten sera of metastasised patients to determine the overall prevalence of methylation changes in breast cancer. As a next step the inventors analysed the 33 genes that were positive in the metastasised patients, in the pre-treatment sera of 26 patients with primary breast cancer and in ten benign controls in order to identify the most important genes for further analysis. Eventually the inventors came up with five genes (ESR1, APC, HSD17B4, HIC1 and RASSF1A), which were primarily analysed in a group of 24 patients (training set). To confirm the significance of this result the inventors tested these genes in an independent set of 62 patients (test set). In order to apply the strictest criteria for testing the potential of a prognostic factor, the inventors investigated these markers in women, who had not undergone adjuvant systemic treatment. DNA methylation of APC and RASSF1A in pretherapeutic sera, both frequently methylated and abnormally regulated in human primary breast cancers (Dammann et al.: Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. Cancer Res., 61: 3105-3109, 2001; Virmani et al.: Aberrant methylation of the adenomatous polyposis coli (APC) gene promoter 1A in breast and lung carcinomas. Clin. Cancer Res., 7: 1998-2004, 2001), turned out to be a strong independent prognostic parameter. These genes are involved in pathways counteracting metastasis: mediation of intercellular adhesion, stabilisation of the cytoskeleton, regulation of the cell cycle and apoptosis (Fearnhead et al.: The ABC of APC. Hum. Mol. Genet., 10: 721-733, 2001; Dammann et al.: Epigenetic inactivation of the Ras-association domain family 1 (RASSF1A) gene and its function in human carcinogenesis. Histol. Histopathol., 18: 665-677, 2003). Methylated DNA in patients' pretherapeutic serum coding for these two genes reflects poor prognosis. The source of the tumour-specific DNA and its definite role in metastasis remains elusive. Circulating tumour-specific altered genetic information may serve as a surrogate marker for circulating tumour cells that ultimately cause distant metastases. An alternative, but equally attractive, hypothesis is that circulating altered DNA per se may cause de novo development of tumour cells in organs known to harbour breast cancer metastases. This so-called "Hypothesis of Genometastasis" suggests that malignant transformation might develop as a result of transfection of susceptible cells in distant target organs with dominant oncogenes that circulate in the plasma and are derived from the primary tumour. Interestingly, irrespective of the source of DNA in the serum, it is noteworthy that some genes provide prognostic information when methylated in patients' sera, whereas genes like HIC1, which is

methylated in about 40% and 90% of primary and metastasised breast cancer patients, respectively, but in only 10% of healthy individuals, are not at all a prognostic parameter.

Irrespective of the mechanistic role of methylated DNA with regards to metastasis in breast cancer patients, these epigenetic changes in serum have several advantages as indicators of poor prognosis as compared to currently used or studied prognostic parameters: DNA in serum is stable and can be analysed by a high-throughput method like MethyLight. Compared to bone marrow aspiration, a simple blood draw (which can be repeated any time throughout the follow-up period) is sufficient. The more screening mammographies are performed, the more small cancers are treated and after histopathological examination no tumour material will remain to perform RNA- and/or protein-based assays for risk evaluation. This application therefore demonstrates a useful and easy approach for risk assessment of breast cancer patients.

Table 1

Gene	Genomic sequence	Bisulphite sequence
HIC1 NM_006497	1	6, 7, 16 and 17
HSD17B4 NM_000414	2	8, 9, 18 and 19
APC NM_000038	3	10, 11, 20 and 21
ESR1 NM_000125	4	12, 13, 22 and 23
RASSF1A NM_170715	5	14, 15, 24 and 25

Table 2. Characteristics of training and test sets.

Characteristics	Training Set (N=24)	Test Set (N=62)	P Value*
	percent		
Size of tumour			0.024
T1	62.5	79	
T2	37.5	13	
T3 + T4	0	7	
Histologic type			n.s
Invasive ductal	67	63	
Invasive lobular	8	13	
Others	25	24	

Tumor grade			n.s
1	46	44	
2	33	39	
3	17	10	
Lymph node metastases			n.s.
No	75	65	
Yes	12.5	11	
Unknown	12.5	24	
Menopausal status			n.s.
Premenopausal	33	16	
Postmenopausal	67	84	
Estrogen-receptor status			n.s.
Positive	54	40	
Negative	42	45	
Progensteron-receptor status			n.s.
Positive	58	45	
Negative	38	40	
Hormone-receptor status			n.s.
Positive	63	50	
Negative	33	36	

*P values for the comparison of numbers of patients were calculated by means of the Chi^2 test. n.s., not significant; Median age: training set (54.2 years; 37.6-83.2), test set (65.7 years; 28.2-86.2), P = 0.052 ; Follow-up: training set (8.0 years; 1 month to 12.2. years), test set (3.1 years; 1 month to 11. years) P < 0.001.

Tumour grade was unknown in six cases.
Hormone-receptor status was unknown in ten cases.
Tumour size was unknown in one case.

Table 3. Frequency of methylated serum DNA in the gene evaluation set.

Gene	Healthy Controls (N = 10)	Primary Breast Cancer (N = 26)	Recurrent Breast Cancer (N = 10)
		percent positive	
ESR1	0	27	70
APC	0	23	80
HSD17B4	0	12	30
CDH13	0	8	40
ESR2	0	4	20
MGMT	0	4	10
SYK	0	4	10
HIC1	10	39	90
RASSF1A	10	23	80
GSTP1	10	12	60
MYOD1	20	27	80
CDH1	20	20	90

PTGS2	30	39	100
PGR	30	46	80
CALCA	40	50	60
HLAG	60	69	100
BLT1	60	85	100
ARHI	100	100	100
MLLT7	100	100	100
TFF1	100	100	100
SOCS2	0	0	40
SOCS1	0	0	30
TERT	0	0	30
DAPK1	0	0	30
TIMP3	0	0	20
BRCA1	0	0	20
GSTM3	0	0	20
MT3	0	0	20
TWIST	0	0	10
MLH1	0	0	10
CYP1B1	0	0	10
TITF1	0	0	10
FGF18	0	0	10
CDKN2A	n.d.	n.d.	0
HSPA2	n.d.	n.d.	0
PPP1R13 B	n.d.	n.d.	0
TP53BP2	n.d.	n.d.	0
REV3L	n.d.	n.d.	0
IGFB2	n.d.	n.d.	0

n.d., not done

Table 4. Univariate analysis of methylation status in *training* and *test sets*.

Genes	Training Set (N=24) P Value	Test Set (N=62) P Value
ESR1	0.018	0.555
APC	0.002	0.307
HSD17B4	0.056	
HIC1	0.796	
RASSF1A	0.042	0.014
RASSF1A/APC	<0.001	0.007
ESR1/APC	0.001	0.951
ESR1/RASSF1A	0.032	0.138

*P values for each variable were calculated by means of the log rank test.

Table 5. Frequency of methylated genes according to clinicopathological features.

Characteristics	No. of Patients	ESR1	APC	RASSF1 A	RASSF1 A and/or APC
		% positive			
Size of tumour					
T1	64	14	11	9	19
T2	17	12	12	19	31
T3 + T4	4	75	25	50	50
Histologic type					
Invasive ductal	55	18	15	11	22
Invasive lobular	10	20	0	30	30
Others	21	10	10	10	20
Tumor grade					
1	38	11	11	13	21
2	32	19	16	16	31
3	10	30	10	11	11
Lymph node metasta-ses					
No	58	12	9	9	18
Yes	10	20	30	20	40
Unknown	18	28	11	22	28
Menopausal status					
Premenopausal	18	28	11	11	22
Postmenopausal	68	13	12	13	22
Estrogen-receptor					
Positive	38	16	11	16	21
Negative	38	16	16	11	27
Progensterone-receptor					
Positive	42	14	5	14	18
Negative	34	18	24	12	33
Hormone-receptor status					
Positive	46	15	9	15	20
Negative	30	17	20	10	31

Tumour grade was unknown in six cases. Hormone-receptor status was unknown in ten cases. Tumour size was unknown in one case. DNA methylation of RASSF1A for one case was missing. Chi^2 Pearson: Tumour size – ESR1 (P = 0.005); Tumour size – RASSF1A (P = 0.049); Progesterone-receptor – APC (P = 0.036); Median age – RASSF1A methylated (79.0 yrs.; 49.6 to 86.2), RASSF1A unmethylated (59.4 yrs.; 28.2 to 82.3.) P = 0,009

Table 6. Results of univariate analysis.

Variable	No. of Patients Who Died/Total No.	Relative Risk of Death (95% CI)	P Value
Size of tumour			0.018
T1	10/64		
T2	5/17	2.2 (0.6 - 7.8)	

T3 + T4	2/4	5.4 (0.7 - 42.9)	
Histologic type			0.296
Invasive ductal	13/55		
Invasive lobular	1/10	0.4 (0 - 3.1)	
Others	3/21	0.5 (0.1 - 2.1)	
Tumor grade			0.310
1	6/38		
2	9/32	2.1 (0.7 - 6.7)	
3	2/10	1.3 (0.2 - 7.9)	
Lymph node metastases			0.005
No	7/58		
Yes	5/10	7.3 (1.7 - 31.7)	
Unknown	5/18	2.8 (0.8 - 10.3)	
Menopausal status			0.062
Premenopausal	1/18		
Postmenopausal	16/68	5.2 (0.6 -42.4)	
Estrogen-receptor			0.369
Positive	10/38	1.9 (0.6 - 5.9)	
Negative	6/38		
Progensterone-receptor			0.766
Positive	9/42	1.1 (0.3 - 3.2)	
Negative	7/34		
Hormone-receptor status			0.799
Positive	10/46	1.1 (0.4 - 3.5)	
Negative	6/30		
ESR1 methylation			0.370
Unmethylated	13/72		
Methylated	4/14	1.8 (0.5 - 6.7)	
APC methylation			0.001
Unmethylated	12/76		
Methylated	5/10	5.3 (1.3 - 21.3)	
RASSF1A methylation			0.001
Unmethylated	11/74		
Methylated	6/11	6.9 (1.8 - 26.5)	
RASSF1/APC methylation			<0.001
Unmethylated	7/66		
Methylated	10/19	9.5 (2.9 - 31.4)	

Table 7. Multivariate Analysis.

Variable	Relative Risk of Death (95% CI)	P Value
Size of tumour		0.19
T2 (vs. T1)	2.7 (0.8-9.3)	
T3 + T4 (vs. T1)	2.9 (0.4-20.5)	
Lymph node metastases		0.039
Yes (vs. no lymph node metastases)	3.9 (1.1-13.9)	

Unknown (vs. no lymph node metastases)	5.2 (1.2-22.4)	
Age	1.0 (1.0-1.1)	0.06
RASSF1A and/or APC methylated (vs. unmethylated)	5.7 (1.9-16.9)	0.002

Table 8 Sequences of the primers and probes

HUGO Gene Nomenclature	Forward Primer Sequence	Reverse Primer Sequence	Probe Oligo Sequence
ACTB	TGGTGATGGAGGAGGTTTAGTAAGT	AACCAATAAAACCTACTCCTCCCTTAA	6FAM-ACCACCACCCAACACACAATAACAAACACA-BHQ-1
APC	GAACCAAAACGCTCCCCAT	TTATATGTCGGTTACGTGCGTTTATAT	6FAM-CCCGTCGAAAACCCGCCGATTA-BHQ-1
ARHI	GCGTAAGCGGGAATTTATGTTTGT	CCGGCGATTTTATATTCCGACTT	6FAM-CGCACAAAAACGAAAATACGAAAACGCAAA-BHQ-1
BLT1	GCGTTGGTTTTATCGGAAGG	AAACCGTAATTCCCGCTCG	6FAM-GACTCCGCCCAACTTCGCCAAAA-BHQ-1
BRCA1.	GAGAGGTTGTGTGTTTAGCGGGTAGTT	CGGGCAATCGCAATTTTAAT	6FAM-CCGGCGTTTTCCGTTACCACGA-BHQ-1
CALCA	GTTTTGGAAGTATGAGGGTGACG	TTCCCGCGCGCTATAAATCG	6FAM-ATTCCGCCAATACACAACAACCAATAAACG-BHQ-1
CDH1	AATTTTAGGTTAGAGGGTTATCGCGT	TCCCCAAAACGAAACTAACGAC	6FAM-CGCCCACCCGACCTCGCAT-BHQ-1
CDH13	AATTTCGTTCGTTTTGTGCGT	CTACCCGTACCGAACGATCC	6FAM-AACGCAAAACGGCCCGACA-BHQ-1
CDKN2A	TGGAGTTTTTCGGTTGATTGGTT	AACAACGCCCGCACCTCCT	6FAM-ACCCGACCCCGAACGCG-BHQ-1
CYP1B1	GTGCGTTTGGACGGGAGTT	AACGCGACCTAACAAAACGAA	6FAM-CGCCGCACCACCAAACCGCTT-BHQ-1
DAPK1	TCGTCGTTTCGTTTCGGTTAGTT	TCCCTCCGAAACGCTATCG	6FAM-CGACCATAAACGCCAACGCCG-BHQ-1
ESR1	GGCGTTCGTTTTGGGATTG	GCCGACACGCGAACTCTAA,	6FAM-CGATAAAACCGAACGACCCGACGA-BHQ-1
ESR2	TTTGAAATTTGTAGGGCGAAGAGTAG	ACCCGTCGCAACTCGAATAA	6FAM-CCGACCCAACGCTCGCCG-BHQ-1
FGF18	ATCTCCTCCGCGGTCTCT	TCGCGCGGTAGAAAACGTTT	6FAM-CGACCGTACGCATCGCGGC-BHQ-1
GSTM3	GCG CGA ACG CCC TAA CT	AAC GTC GGT ATT AGT CGC GTT T	6FAM-CCC CGT TCT CCG TCC CTT ACC TCC-BHQ-1
GSTP1	GTCGGGCGTCGTGATTTAGTATTG	AAACTACGACGACGAAACTCCAA	6FAM-AAACCTCGGACCTCCGAACCTTATAAAA-BHQ-1
HIC1	GTTAGGCGGTTAGGGCGTC	CCGAACGCCTCCATCGTAT	6FAM-CAACATCGTCTACCCAACACACTCTCTACG-BHQ-1
HLA-G	CAC CCC CAT ATA CGC GCT AA	GGT CGT TAC GTT TCG GGT AGT TTA	6FAM-CGC GCT CAC ACG CTC AAA AAC CT-BHQ1
HSD17B4	TATCGTTGAGGTTCGACGGG	TCCAACCTTCGCATACTCACC	6FAM-CCCGGCGCCGATAACCAATACCA-BHQ-1
HSPA2	CAC GAA CAC TAC CAA CAA CTC	GGG AGC GGA TTG GGT TTG	6FAM-CCG CGC CCA ATT CCC GAT TCT-BHQ1

Gene			
IGFBP2	AACT	CGG GAA GAG TAG GGA ATT TTT AGA GT	6FAM-ACG CCC GCT CGC CCA CCT-BHQ1
MGMT	CTC GCG CCG ACA AAT AAA TAC GCGGTTTCGACGTTCGTAGGT	CACTCTTCCGAAAACGAAACG	6FAM-CGCAAACGATACGCACCGCGA-BHQ-1
MLH1	AGGAAGAGCGGATAGCGATTT	TCTTCGTCCCTCCCTAAAACG	6FAM-CCCGCTACCTAAAAAATATACGCTTACGCG-BHQ-1
MLLT7	CCT CAC GAT ACC TCC CCT CAA CGA TAA ACG AAC TTC TCC AAA CAA	TTA GGG ATT AGC GTT TTG GGA TT	6FAM-AAA CAC ATT CCT ACC AAT CTT CAA AAA ATC GCG-BHQ1
MT3		GCG CGG TGC GTA GGG	6FAM-AAA CGC GCG ACT TAA CTA ATA ACA ACA AAT AAC GA-BHQ-1
MYOD1	GAGCGGCGTAGTTAGCG	TCCGACACGCCCTTCC	6FAM-CTCCAACACCCGACTACTATATCCCGAAA-BHQ-1
PGR	TTATAATTCGGAGGCGGGTTAGTGTTT	TCGAACTTCTACTAACTCCGTACTACGA	6FAM-ATCATCTCCGAAAATCTCAAATCCAATAATAC G-BHQ-1
PPP1R13B	CCT CAC CCA CCG ACA TCA TC	TCG GAG CGG TGG GTA TAG TTC	6FAM-AAA AAT CCG CGA CGC CCT CGA-BHQ-1
PTGS2	CGGAAGCGTTCGGGTAAAG	AATTCCACCGCCCCAAAC	6FAM-TTTCCGCCAAATATCTTTCTTCTTCGCA-BHQ-1
RASSF1A	ATTGAGTTGCGGGGAGTTGGT	ACACGCTCCAACGAATACG	6FAM-CCCTTCCCAACGCGCCCA-BHQ-1
REV3L	CGA ACG CAA CCG ACC CT	TAT TTT TCG TAT CGT TTT CGG GTT A	6FAM-CTC AAA TAA CGC CGC GAC TCC GC-BHQ-1
SOCS1	GCGTCGAGTTCGTGGGTATTT	CCGAAACCATCTTCACGCTAA	6FAM-ACAATTCCGCTAACGACTATCGGCA-BHQ-1
SOCS2	TCC CTT CCC CGC CAT T	TTG TTT TTG TCG CGG TGA TTT	6FAM-CCG AAA AAC TCA AAA CAC CGC AAA ATC AT-BHQ1
SYK	GGGCGCGATATTGGGAG	GCGACTCTTCCTCATTTAAACAAC	6FAM-CCTTAACGCGCCGAACAAACG-BHQ-1
TERT	GGATTCGCGGGTATAGACGTT	CGAAATCCGCGCGAAA	6FAM-CCCAATCCCTCGCCACGTAAAA-BHQ-1
TFF1	TAAGGTTACGGTGGTTATTTCGTGA	ACCTTAATCCAAATCCTACTCACTATCTA AAA	6FAM-CCCTCCCGCCAAATAATAAATACTATACTACTAC AAAA-BHQ-1
TIMP3	GGCGTCGGGAGGTTAAGGTTGTT	CTCTCCAAAATTACCGTACGCG	6FAM-AACTCGCTCGCCCGCCGAA-BHQ-1
TITF1	CGA AAT AAA CCG AAT CCT CCT TAA	TGT TTT GTT GTT TTA GCG TTT ACG T	6FAM-CTC GCG TTT ATT TTA ACC CGA CGC CA-BHQ-1
TP53BP2	ACC CCC TAA CGC GAC TTT ATC	GTT CGA TTC GGG ATT AGT TGG T	6FAM-CGC TCG TAA CGA TCG AAA CTC CCT CCT-BHQ-1
TWIST	GTAGCGGCGGCGAACGT	AAACGCAACGAATCATAACCAAC	6FAM-CCAACGCACCCAATCGCTAAACGA-BHQ-1

Epigenomics AG
E30120

Claims

1.   A method for determining the prognosis of a subject with a cell proliferative disorder of the breast tissues, said method comprising analysing the methylation pattern of a target nucleic acid comprising one or a combination of the genes taken from the group consisting of ESR1, APC, HSD174B4, HIC1 and RASSF1A and/or their regulatory regions by contacting at least one of said target nucleic acids in a biological sample obtained from said subject with at least one reagent, or series of reagents that distinguishes between methylated and non-methylated CpG dinucleotides.

2.   A method for selecting a treatment of a cell proliferative disorder of the breast tissues, said method comprising:

a)   determining the prognosis of a subject according to Claim 1, and

b)   selecting a suitable treatment according to said prognosis.

3.   A method for determining the phenotype of a subject with a breast cell proliferative disorder comprising

a) obtaining a biological sample containing genomic DNA from said subject,

b) analysing the methylation pattern of one or more target nucleic acids comprising one or a combination of the genes taken from the group consisting of ESR1, APC, HSD174B4, HIC1 and RASSF1A and/or their regulatory regions by contacting at least one of said target nucleic acids in the biological sample obtained from said subject with at least one reagent, or series of reagents that distinguishes between methylated and non-methylated CpG dinucleotides, and

c) determining the phenotype of the individual by comparison to two known phenotypes, a first phenotype characterised by hypermethylation of the target nucleic acid and poor progno-sis as relative to a second phenotype characterised by hypomethylation of the analysed target nucleic acid and positive prognosis.

4.   A method according to claims 1 to 3 wherein said prognosis is the life expectancy of said subject.

5.   A method according to any one of claims 1 to 4 wherein said target nucleic acid com-prises the gene APC and/or its regulatory regions.

6.   A method according to any one of claims 1 to 4 wherein said target nucleic acid comprises the gene RASSF1A and/or its regulatory regions.

7.   A method according to any one of claims 1 to 4 wherein said target nucleic acids comprise the genes APC and RASSF1A and/or their regulatory regions.

8.   A method according to any one of claims 1 to 7, wherein said target nucleic acid or acids comprise essentially one or more sequences from the group consisting of SEQ ID NOs: 1 to 5 and sequences complementary thereto.

9.   A method according to claim 5 wherein the sequence of said target nucleic acid comprises essentially of SEQ ID NO: 3.

10.   A method according to claim 6 wherein the sequence of said target nucleic acid comprises essentially of SEQ ID NO: 5.

11.   A method according to Claim 7, wherein said target nucleic acid or acids comprise essentially of SEQ ID NOs: 3 and 5.

12.   A method according to Claims 1 to 11, wherein said cell proliferative disorder of the breast tissue is selected from the group consisting of ductal carcinoma in situ, lobular carcinoma, colloid carcinoma, tubular carcinoma, medullary carcinoma, metaplastic carcinoma, intraductal carcinoma in situ, lobular carcinoma in situ and papillary carcinoma in situ.

13.   A nucleic acid molecule consisting essentially of a sequence at least 18 bases in length according to one of the sequences taken from the group consisting of SEQ ID NOs: 6 to 25 and sequences complementary thereto.

14.   An oligomer, in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer, said oligomer consisting essentially of at least one base sequence having a length of at least 10 nucleotides which hybridises to or is identical to one of the nucleic acid sequences according to SEQ ID NOs: 6 to 25.

15.   The oligomer as recited in any one of Claims 13 or 14 , wherein the base sequence includes at least one CpG dinucleotide.

16.   A set of oligomers, comprising at least two oligomers according to any of claims 13 or 14.

17.   A set of oligonucleotides as recited in Claim 16, characterised in that at least one oligonucleotide is bound to a solid phase.

18.   A set of at least two oligonucleotides as recited in any of claims 14 or 15, which is used as primer oligonucleotides for the amplification of nucleic acid sequences comprising one of SEQ ID NOs: 6 to 25 and sequences complementary thereto.

19.   Use of a set of oligonucleotides comprising at least two of the oligomers according to any of claims 16 to 18 for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) within the sequences taken from the group SEQ ID NOs: 1 to 5 and sequences complementary thereto.

20.   A method for manufacturing an arrangement of different oligomers (array) fixed to a carrier material for predicting the responsiveness of a subject with a cell proliferative disorder of the breast tissues by analysis of the methylation state of any of the CpG dinucleotides of the group SEQ ID NOs 1 to 5 wherein at least one oligomer according to any of the claims 14 or 15 is coupled to a solid phase.

21.   An arrangement of different oligomers (array) obtainable according to claim 20.

22.   An array of different oligonucleotide- and/or PNA-oligomer sequences as recited in Claim 21, characterised in that said oligonucleotides are arranged on a plane solid phase in the form of a rectangular or hexagonal lattice.

23.   The array as recited in any of the Claims 21 or 22, characterised in that the solid phase surface is composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold.

24.   A DNA- and/or PNA-array for predicting breast cell proliferative disorders' response by analysis of the methylation state of any of the CpG dinucleotides of the group SEQ ID NOs: 1 to 5 comprising at least one nucleic acid according to any of the claims 14 to 18.

25.   A method according to any one of Claims 1 to 4 comprising the following steps:

a) obtaining a biological sample containing genomic DNA,

b) extracting the genomic DNA,

c) converting cytosine bases in the genomic DNA sample which are unmethylated at the 5-position, to uracil or another base which is dissimilar to cytosine in terms of base pairing behaviour,

d) amplifying at least one fragment of the pretreated genomic DNA, wherein said fragments comprise one or more sequences selected from the group consisting of SEQ ID NOs: 6 to 25 and sequences complementary thereto, and

e) determining the methylation status of one or more genomic CpG dinucleotides by analysis of the amplificate nucleic acids.

26.   The method as recited in Claim 25, characterised in that Step e) is carried out by means of hybridisation of at least one oligonucleotide according to Claims 14 or 15.

27.   The method as recited in Claim 25, characterised in that Step e) is carried out by means of hybridisation of at least one oligonucleotide according to Claims 14 or 15 and extension of said hybridised oligonucleotide(s) by at least one nucleotide base.

28.   The method as recited in Claim 25, characterised in that Step e) is carried out by means of sequencing.

29.   The method as recited in Claim 25, characterised in that Step d) is carried out using methylation specific primers.

30.   The method as recited in Claim 25, further comprising in step d) the use of at least one nucleic acid molecule or peptide nucleic acid molecule comprising in each case a contiguous sequence at least 9 nucleotides in length that is complementary to, or hybridises under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NOs: 6 to 25, and complements thereof, wherein said nucleic acid molecule or pep-

tide nucleic acid molecule suppresses amplification of the nucleic acid to which it is hybrid-ised.

31.   The method as recited in Claim 25, characterised in that Step e) is carried out by means of a combination of at least two of the methods described in Claims 26 to 30.

32.   The method as recited in Claim 25, characterised in that the treatment is carried out by means of a solution of a bisulfite, hydrogen sulfite or disulfite.

33.   A method according to any one of Claims 1 to 12 comprising the following steps:
a)   obtaining a biological sample containing genomic DNA,
b)   extracting the genomic DNA,
c)   digesting the genomic DNA comprising one or more of the sequences from the group consisting of SEQ ID NOs: 1 to 5 and sequences complementary thereto with one or more methylation sensitive restriction enzymes, and
d)   determining of the DNA fragments generated in the digest of step c).

34.   A method according to claim 33, wherein the DNA digest is amplified prior to step d).

35.   The method as recited in any one of Claims 25 to 32 and 34, characterised in that more than six different fragments having a length of 100 - 200 base pairs are amplified.

36.   The method as recited in any one of Claims 25 to 32, 34 and 35, characterised in that the amplification of several DNA segments is carried out in one reaction vessel.

37.   The method as recited in any one of Claims 25 to 32 and 34 to 36, characterised in that the polymerase is a heat-resistant DNA polymerase.

38.   The method as recited in any one of Claims 25 to 32 and 34 to 37, characterised in that the amplification is carried out by means of the polymerase chain reaction (PCR).

39.   The method as recited in any one of Claims 25 to 32 and 34 to 38, characterised in that the amplificates carry detectable labels.

40. · ·The method according to Claim 39, wherein said labels are fluorescence labels, radio-nuclides and/or detachable molecule fragments having a typical mass which can be detected in a mass spectrometer.

41. The method as recited in any one of Claims 25 to 32 and 34 to 40, characterised in that the amplificates or fragments of the amplificates are detected in the mass spectrometer.

42. The method as recited in any one of the Claims 40 or 41, characterised in that the produced fragments have a single positive or negative net charge for better detectability in the mass spectrometer.

43. The method as recited in any one of Claims 40 to 42, characterised in that detection is carried out and visualised by means of matrix assisted laser desorption/ionisation mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI).

44. The method as recited in any one of the Claim 1 to 25 or one of the Claims 25 to 43, characterised in that the genomic DNA is obtained from cells or cellular components which contain DNA, sources of DNA comprising, for example, cell lines, histological slides, biopsies, tissue embedded in paraffin, breast tissues, blood, plasma, lymphatic fluid, lymphatic tissue, duct cells, ductal lavage fluid, nipple aspiration fluid and combinations thereof.

45. A kit comprising a bisulfite (= disulfite, hydrogen sulfite) reagent as well as oligonucleotides and/or PNA-oligomers according to any one of the Claims 14 to 15.

46. A kit according to claim 45, further comprising standard reagents for performing a methylation assay from the group consisting of MS-SNuPE, MSP, Methyl light, Heavy Methyl, nucleic acid sequencing and combinations thereof.

47. The use of a method according to any one of claims 1 to 12 and 25 to 44, a nucleic acid according to claim 13, of an oligonucleotide or PNA-oligomer or a set thereof according to any one of claims 14 to 18, of a kit according to Claim 45 or 46, of an array according to any one of claims 21 to 24 or of a method of manufacturing an array according to claim 20 in the prognosis, diagnosis, treatment, characterisation, classification and/or differentiation of breast cell proliferative disorders. ·

Epigenomics AG
E30120

Abstract

The present invention relates to modified and genomic sequences, to oligonucleotides and/or PNA-oligomers for detecting the cytosine methylation state of genomic DNA, as well as to a method for the prognosis and treatment of a cell proliferative disorder of the breast tissues.
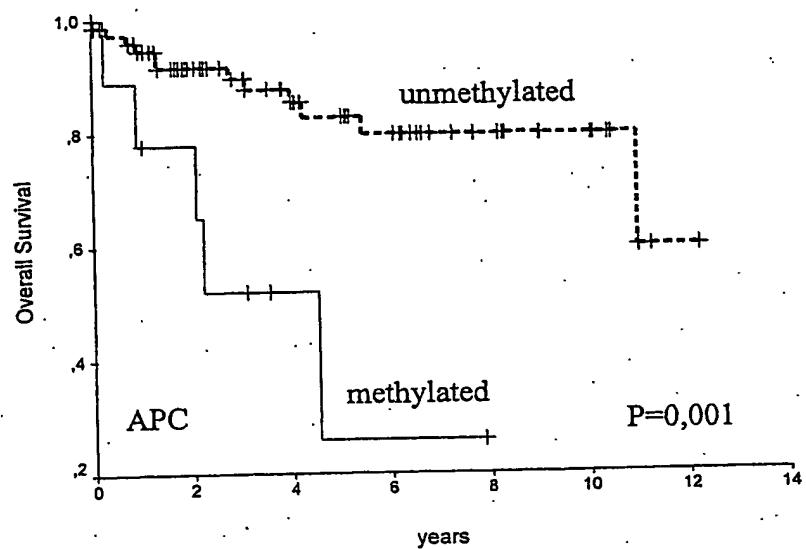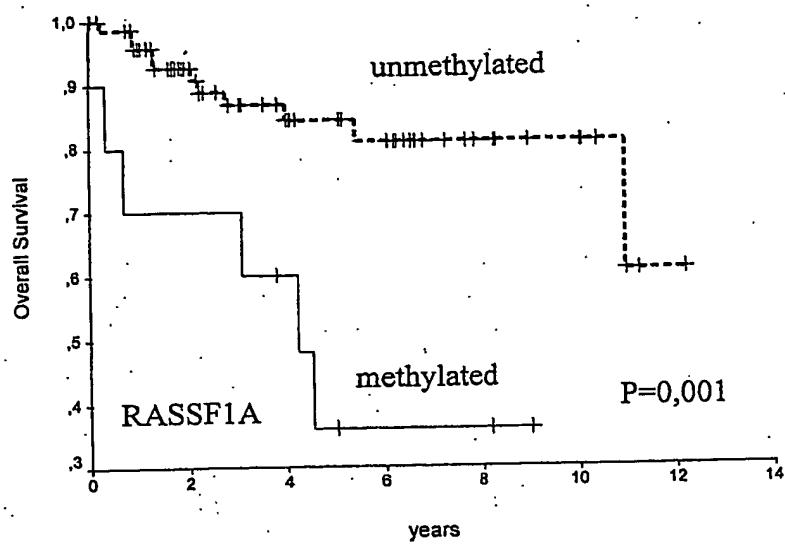
Figure 1

Figure 1

Figure 2

Figure 3


Figure 3
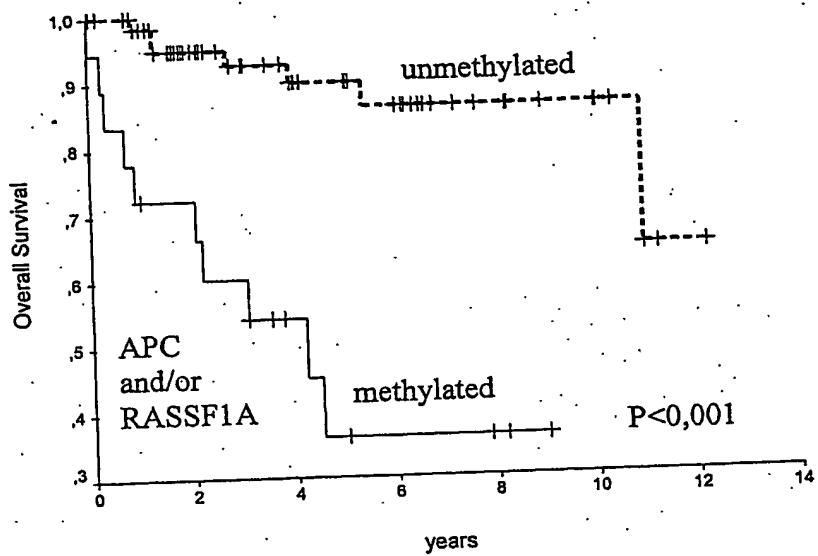
Sequence listing

<110> Epigenomics AG

<120> Prognostic and diagnostic markers for breast cell proliferative
     disorders.

<160> 25

<210> 1
<211> 3501
<212> DNA
<213> Homo Sapiens

<400> 1

```
gcggggctgg caggggcgct gccctggcac agctcggggc ctggcagcgg cgggtggggc     60
atcggctaag agctgccacc gccgcgggga ggggagcccg gcccgccggg accgcaggta    120
acgggccgcg gggccccgcg ggccaggagg ggaacggggt cgggcgggcg agcagcgggc    180
aggggagctc agggctcggc tccgggctct gccgccggat ttggggggcg cgaggaagag    240
ctgcgagccg agggcctggg gccggcgcac tcctcccgcc ctgtctgcag ttggaaaact    300
tttccccaag tttggggcgg cggagttccg ggggagaagg ggccggggga gccgcggagg    360
gaggcgccgg gcccgcgcgt gtagggccca ggccgaggcc gggacgcggg tggggcgcag    420
gcccgggtca gggccgcagc cggctgtgcg ccgtgccgc ccggggcgct gccccctccc    480
tccctggga gctgcgtggc tccccctcc ccccacctg cttcctgcct cagcctcctg    540
ccccgatata acgccctccc cgcgccgggc ccggccttcg cgctctgccc gccacggcag    600
ccgctgcctc cgctccccgc gcggccgccg cccgggcccc gaccgagggt tgacagcccc    660
cggccagggc ggcgccaggg cgggcaccgc gctcccctcc tccgtatcac ttcccccaac    720
tggggcaact tctcccgagg cgggaggcgc tggttcctcg gctcccttc tccctacttg    780
ggtaaagttc tccgccctga atgacttttc ctgaagcgga cattttactt aaatcgggta    840
actgtctcca aaagggtcac tgcgcctgaa cagttttctt ctcggaagcc ccagcaccca    900
gccaggtgcc ctggggcgtg caggccgccc tggcctcccc tccaccggcg gccgctcacc    960
tcctgctcct tctcctggtc cgggcggggcc ggcctgggct cccactccag agggcagccg   1020
gtccttcgcc ggtgcccagg ccgcagggct gatgccccg ctcagctgag ggaaggggaa   1080
gtggagggga gaagtgccgg gctggggcca ggcggccagg gcgccgcacg gctctcaccc   1140
ggccggtgtg tgtccccgca ggagagtgtg ctgggcagac gatgctggac acgatggagg   1200
cgcccggcca ctccaggcag ctgctgctgc agctcaacaa ccagcgcacc aagggcttct   1260
tgtgcgacgt gatcatcgtg gtgcagaacg ccctcttccg cgcgcacaag aacgtgctgg   1320
cggccagcag cgcctacctc aagtccctgg tggtgcatga caacctgctc aacctggacc   1380
atgacatggt gagcccggcc gtgttccgcc tggtgctgga cttcatctac accggccgcc   1440
tggctgacgg cgcagaggcg gctgcggccg cggccgtggc cccggggggct gagccgagcc   1500
tgggcgccgt gctggccgcc gccagctacc tgcagatccc cgacctcgtg gcgctgtgca   1560
agaaacgcct caagcgccac ggcaagtact gccacctgcg gggcggcggc ggcggcggcg   1620
gcggctacgc gccctatggt cggccggggc ggggcctgcg ggccgccacg ccggtcatcc   1680
aggcctgcta cccgtcccca gtcgggcctc ccgagctgta cgcgtcggga cccggcccgg   1740
gcccagaggc cgcggtcaac acgcactgcg gctcccctct ttgtggcctg gacctgtcca   1800
ccgccgcact ctgtgcctcg gagcgccgct gctccctct ggctgagcgc gagctgcccc   1860
agaagagccc gccgggctcc gcggcgccag agccgccgcta caaggagccg cctctcgccc   1920
cgcgcccgga cagccctccc agcgccggcc ccgccgccgg ggccgcaccg ccttccgacc   1980
tgccgtcgct gccgccgctg cccttccaga agctggagga ggccgcaccg ccttccgacc   2040
catttcgcgg cggcagcggc agcccgggac ccgagccccc cggccgcccc gacgggccta   2100
gtctcctcta tcgctggatg aagcacgagc cgggcctggg tggtgggggac gcggccgtct   2160
gccgggagcg cggctccccc agcgagcgct gcgaagagcg taccctggc agcctggacg   2220
cgcccggggg gcccccgctc ggcctggcgc cgccgccgcg cagcgaggag accggtagca   2280
ggcccggcgc gggcggcgac ggcgacgact acaagagcag cccatgcccg cacctggcct   2340
gcgaggaccc cagcccgcct ggcggccacc tcgagggcta cattccgtgc ggcaagggct   2400
atggcgagcc cgagagcttc ggtgacaacc tgtacgtgtg cattccgtgc ggcaagggct   2460
tccccagctc tgagcagctg aacgcgcacg tggaggctca cgtggaggag gaggaagcgc   2520
tgtacggcag ggccgaggcg ccgaagtgg ccgctggggc cgccggccta gggccccctt   2580
ttggaggcgg cggggacaag gtcgccgggg ctccgggtgg cctgggagag ctgctgcggc   2640
cctaccgctg cgcgtcgtgc gacaagagct acaaggaccc ggccacgctg cggcagcacg   2700
agaagacgca ctggctgacc cggccctacc catgcaccat ctgcgggaag aagttcacgc   2760
agcgtgggac catgacgcgc cacatgcgca gccacctggg cctcaagccc ttcgcgtgcg   2820
```

```
acgcgtgcgg catgcggttc acgcgccagt accgcctcac ggagcacatg cgcatccact    2880
cggacgagaa gccctacgag tgccaggtgt gcggcggcaa gttcgcacag caacgcaacc    2940
tcatcagcca catgaagatg cacgccgtgg ggggcgcggc cggcgcggcc ggggcgctgg    3000
cggacttggg ggggctcccc ggctcgcctca cggccgagca gctgagcctg aagcagcagg    3060
ccgagggcgt ctttgctgtg gctcgcctca cggccgagca gctgagcctg aagcagcagg    3120
acaaggcggc cgcggccgag ctgctggcgc agaccacgca cttcctgcac gaccccaagg    3180
tggcgctgga gagcctctac ccgctggcca agttcacggc cgagctgggc ctcagccccg    3240
acaaggcggc cgaggtgctg agccagggcg ctcacctggc ggccgggccc gacggccgga    3300
ccatcgaccg tttctctccc acctagagcg cccctcgcca gcccgctctg tcgctgctgc    3360
gcggccctgg cccgcacccc agggagcggc ggggcggcg cgcagggccc actgtgcccg      3420
ggacaaccgc agcgtcgcca cagtggcggc tccacctctc ggcggcctca cctggcctca    3480
ctgcttcgtg ccttagctcg g                                            3501
```

<210> 2
<211> 2501
<212> DNA
<213> Homo Sapiens

<400> 2

```
tttccatagt gtaaatgtgt tcccaccact ctctggagta atcctactta aaaccgtttt     60
cagcacaaaa ttcaaacatc taaacatgat cttgctggct ttgctttgt ggctttaccc    120
tctttctccc caaacctagc tagtgtttgt gctgcctgta atgcccttct ttctttgcag    180
gggtcgccac tttaggtcct ggtcctcctt cagaaagttt ttcctctttc tccccagcgg    240
ggataggtc tgtttatttt gacaccatta gctcacttac acacattggt cacaagtcta     300
ggctgcaccg ttattgaaag tttaccatct gactctgagt agcttgagga tcctatcaaa    360
actcaggaga tgctcagtaa atgttgattg aactatgact gttctcaaca tacaaacgca    420
agatcattta ggaacacttg tcaaaatgtt tttgcccctt gagattctat tttgggaggt    480
aagcagtggg ggtccaggac tctgcatttt gacagtcccc tgatgtttgc atgtagaagt    540
gcagggatta ttacactgac aaatctttac catccctaag ggggactttc cttcccaggg    600
gctatctctg gaagcccctc aaggataggg gccgcatgct gtttctctag gtcagcaact    660
aaacccagaa aacgtttatt gagtgaatga tgaaacgaca ggtgaataga tgaacgcaag    720
gtgtcgagtt aactattctt ctacacaagt cctagcagct cccattgctt ccagccgcag    780
aaatggcccc tggaaggcaa gtcttccagc gagtggagtc actcttaact acatttccca    840
ggattccaag ggagccgcgc gctctgcgct catcttccta ccagaaatcg gcaagtcact    900
gaccctcgtc ccgcccccgc cattccccgc ctcctcctgt cccgcagtcg gcgtccagcg    960
gctctgcttg ttcgtgtgtg tgtcgttgca ggccttattc atgggctcac cgctgaggtt   1020
cgacgggcgg gtggtactgg tcaccggcgc ggggcaggt gagcatgcga aggttggagg    1080
ccgcgcccct tgctgaggcg cagctggctg ctcttttcgg gccggcatac gcgcgcagcc   1140
gcagctgagg tcaccccgct gaggtggtgg ggaggggaat ggttattctt gaggcaccgc   1200
atctcttgag gaggaaagag ccggaaacac ctggtctctc aagcaggtac agcccgcttc   1260
tccccagcac cccggtgtgg gcttcccaag gtcctgcctg agaggagagg ccaggctggg   1320
ctgctgattg caaaactggg tgaaagttct ccctgaccct tatctgtggg catcgattgt   1380
tactcttcct gcaattaact ctcttagatc tttgcctagt cttttaaagg actgaaaagc   1440
cgcgaggggc ggggctgga attcgccccc tgaagcgcag agatgtcagc tcctgaaaag   1500
tcattcggtc gttcagtgtt tgtttccctc tgtcgtaaga ttttaagttc gtgagaggac   1560
cttctttaaa gagggcgtct gataagagcc cttccccgtt ggagtttgta tgcttagcaa   1620
gtcacaatct gttctcgaaa tccactggag tcttggcaga ggttgtaagc tcaaatgcgc   1680
acagggtca ggcgtatgat ggagaaagaa aatgggagta ggatgggcac atctgaggaa    1740
ctggagagca gagaattccg aagtggaccg gccagtggga aagttgcctg tatttcagga   1800
gcggcaaaat ggaaattgt tatgtgaaat agccccattt ttttaaagtac aaaaaattaa    1860
aacaaaccat tcataccaac atagatgctg tgcagtgaga ttttacatta gtttctcacc   1920
agtgggtgac ctctgtaacc tccaagtgca gggatcttga cattatgcac ctttgattct   1980
ccactggtag taccttatac ctggaaaggc cctaatgcat gaattatttg agttatatat   2040
taaacgttac aaactggaat tctgtcaatt aattcctatg tactttcata tctgtattga   2100
taaagtggct tcttatgctg cctttcagaa aatgctttca gtgttgatga atagccaagt   2160
attttatacc catagctgtc tggttatctc tgcatgggca tgtatttggg tgtagtcata   2220
ccttctaaat gttttagga aaacattttg tttacacttt gcttttattg taaataatgt    2280
attttacaac gcttggtgtt ttaaatcttt tttgacagct cttggataat tttcatgcag   2340
gaggtccagg gattacattc taagacgttt ttgccatcgc taaggagact ttccttttca   2400
ggggctatat ctgaaaatca ttcaaggata gggactgctt cttttgacac cattagcata   2460
cttacacatg gtatgcagta catttacac cagtactcag t                        2501
```

<210> 3
<211> 2470
<212> DNA
<213> Homo Sapiens

<400> 3

```
aaagatgatt aaaagtttaa ttgttcatct gaagagttga tttttttatt cctgtaataa      60
agggtacttt tagcagtctc tgctcatctt gcccatccgg ctcttttgt ggttgtgtaa      120
ggttataact tctgtgtctc agtaaacttg tgcatgccca ttttttttctc tgttactacc      180
ttttctctta ttttgtttta ttattttgat gtaaaattac ctgttaattt tatttgaaat      240
gagaaatttt aaggttcaca ttattcaaat tctgtcagat ccctacctct gtcatatggt      300
ttataatgtg ctgggtattt tcagacctgc ttattaaaaa gatgtaaaac aaaataatga      360
tcactcctgt ggattttttcc tttattttttg agatgtctcc tttggctgca ttacttcttc      420
acccccttgcc cattgatcag aggaggggtc ttaactatgg gtgaaccctа tatcttactg      480
aagaggttat gttacatgta tattttcata atataactta catttacata gtactttat      540
ttttagcata cctttttttta ttaatcctaa taatatcact gtaagttatg ttgaagcaga      600
ttgtaagtgt tcatttacaa attgtgaaat gaattaaaat gaaagggcaa agattaaatc      660
atgaccaggc ctgaaattaa cacacaagac tcaatttttt tcaaccaaag acttttgtag      720
gtgatccctg cctgcaggac tccccttcct cctcagatgt cattggattg taccaggttt      780
actgtagatt ctagccgttg tagaactaac tagatctaag atgagtcccc tgatttcctt      840
tggtagagtc ttccaattgc tgaactccaa tattgtcgtg actagccagt gttacaacct      900
gtctgcctta ttttgtgtaa tggatttcat attacagagg catttttttta atgtcaagat      960
gtttaagtat tgcttaagtg caaactactt aatacttttt agctattaag taattaagat     1020
aggcaggatt ttatttgttc caaaatgatt tgacctaaac taaaaagaga atgtggatct     1080
cctgaatctt acttggttaa tcttaatata actcctagca ttctataatt cttcctaaag     1140
tcctcttacc tggctatctt ttgtatcttc tttgtctctc ctcttctttc ccagtcataa     1200
taactgccag actctgcttc atttctctttt gacagtctct actcctaagg tcatccattc     1260
tctttaggta tcttttggcc tcagtttgag cacagcagat cccaagacca catatgccat     1320
agcataggct attatagtca accttttgaa taaatgtgat tgaactttat gttagtaatt     1380
cttatttacc atcttcctat caaaaaggct taaagtcttc atttaatgct ctccttcatg     1440
tccattttgt taaatgattg ccttttaatg acatcttaga acttcagaac tatttcacca     1500
tggaggatgt gtaagattag cctttttatca aataaaaagt gtgaaatgga atatgtaatc     1560
tcattaatcc attctggctc taaaattctg tgactatcag ataaaattca gaaataaaat     1620
agtattacta atataaataa attttttatca taattatatt tcctaagttt tgcctgtaag     1680
aatgggtaaa atatctttaa aaccttgaag aaattattac ttgatagaaa gtttaatcca     1740
tctgtgagaa ggcaaatgta ttcagacaca actaaagttc tctcttctat tttaatttca     1800
tttatcttga actaagactc cactgtttca tcctcttaga tgctgctact tgaacaatat     1860
tgttttgaga ccaaaaacta gcatattaac acaattcttc ttaaacgtct taagagtttt     1920
gtttcctttta cccctttctt taaaaacaag cagccactaa attttttagt agtgaatttc     1980
aaaatccttt ttaaccttat aggtccaagg gtagccaagg atggctgcag cttcatatga     2040
tcagttgtta aagcaagttg aggcactgaa gatggagaac tcaaatcttc gacaagagct     2100
agaagataat tccaatcatc ttacaaaact ggaaactgag gcatctaata tgaaggtatc     2160
aagactgtga cttttaattg tagtttatcc atttttattc agtattccct cttgtaaact     2220
tgaggtaaga cactttactt aaaagtgtat tttaaattaa gcaataatat gtaaactctt     2280
tcttgcaaaa gttagcattt atattttttaa ataagatata ttgaattcat tcagtgaatc     2340
atataaagaa aataagtgta aaactccaat ggctagttag ttcttagttc tttttaagat     2400
taaagagaag agaccaaata tagcatcact gtactgaggc aaggttttct gtgtagttca     2460
tagaaactag                                                            2470
```

<210> 4
<211> 7001
<212> DNA
<213> Homo Sapiens

<400> 4

```
aatgcaatgg aaaaagagag attgtaaagc tagaaggctt aggaattgcc tcttgattag      60
gtgtggaagg caagggaaaa tcagccctcg aagaagacag tgagatttta atctgggtgg      120
ctggagagac agtgatgctg ggcacagaca cggggaagtt gagaggaaca ccatgtttga      180
gaatggtgac tcatatttga acaagcctgc aatgcccagc agaccgctgg aaaagtgggg      240
ctggagacac attcaacgga ggagccagat caatctttac ccttcttcac ctgagagagc      300
cagtaagtca cggctggaac gtgtgtgtcc agcaggagag ggtagggagg gaagccaaga      360
```

```
gagctgggag cccgagtgaa gttttttgcca aaggcagaag aggaaagtcg gcgtagcaca   420
gtatacttttc ccacccatgc tcaccaagcc cagggacaag gctcaccaag atgagtttgg   480
aagagaatgc tggagagaaa gtggttaaga aaactgcctt tactgaactt cttgggctaa   540
ctttgattgt aagtctctga acaatcaaag cctgtgagga gacagctaac cttcttattc   600
ttcctatgtc aatagtgaac aattgcagat ccccttttcct ttccttctcc tttcccctgt   660
tcctctctcc tccctcccctg aatactcttg cttttttctg ggactggtct agagcatggg   720
tggccattgt tgacctacag gaggcaccac tgtcaccaac aaagggtaac agtctttctt   780
ttcaatattt atttatatcc agtatttatt ttcaatactg actatggaga gagctctcct   840
gtgctcaaac actgcaatac tgggggtctt tcaaagcaca aaaacatata tttgcatgat   900
ggcatcatta acatttttat ggctttctat ttctttttttg tactggtctc aagagccact   960
cataaatctc tcagtaactg catagtgtcc cagggccaga gaccggccac tcctggcatt  1020
gtgattagag tcatttaata tccaaggtgg tgactaatgt ctggcaacaa agcctccatt  1080
gggtgtcatg tgtcctggga ccctgagcgt gggcactcta ggagcacctc agtattgcgt  1140
gttagtacta tggccgagag aatagttgag aaagtggtca agaggtggat ccatgtgaac  1200
gccactggga aatgagagac ctcgttccca atcacggtca gtgcaactcg aaagcctaaa  1260
atcagtttaa aacaaaggta tctacctttta tcttatgttc atatcctagg ctttttaataa  1320
tacgtatttt tcacatgttt acagaaagca gtcaactgag ctattcatgg aaaggtttgt  1380
gggtttggtt aacgaagtgg aggagtatta catttcagct ggaaacacat ccctagaatg  1440
ccaaaacatt tattccaaag tctggtttcc tggtgcaatc ggaggcatgg caatgcctct  1500
gttcagagac tgggggctag ggccagtaag gcatttgatc cacatgtatc ccagaaggct  1560
tttattgtta aattatattc tttcggaaaa accacccatg tcctatttttg taaacttgat  1620
atccatacac ttttgactgg cattctattt tagccgtaag actatgattc acagcaagcc  1680
tgttttttcct cttgcttggg gtggcagcag aaagcatagg gtactttcca gcctccaagg  1740
gtaggggcaa aggggctggg gtttctcctc cccagtacag cttttctctgg ctgtgccaca  1800
ctgctccctg tgagcagaca gcaagtctcc cctcactccc cactgccatt catccagcgc  1860
tgtgcagtag cccagctgcg tgtctgccgg gaggggctgc caagtgccct gcctactggc  1920
tgcttcccga atccctgcca ttccacgcac aaacacatcc acacactctc tctgcctagt  1980
tcacacactg agccactcgc acatgcgagc acattccttc cttccttctc actctctcgg  2040
cccttgactt ctacaagccc atggaacatt tctggaaaga cgttcttgat ccagcagggt  2100
aggcttgttt tgatttctct ctctgtagct ttagcatttt gagaaagcaa cttaccttttc  2160
tggctagtgt ctgtatccta gcaggagat gaggattgct gttctccatg ggggtatgtg  2220
tgtgtctcct tttttctttca ggacttgtag gattctttgt gccatttgca tataatttgg  2280
caggttcaca ttttttaaga gccctatgaa gtgcttttttg catgtgtttt aaaaaggcat  2340
ttgaaaattg aaagtgtgat ttatggaaat taaatcatct gtaaaaaatt gctttggaaa  2400
gtaatgattg ctggccataa agggaaatat ctgcgatgca cctaatgtgt tttttaaccct  2460
ttatttgctg acaatctata gtcattaatg ctaaactcga ttttggcttc agctacattt  2520
gcatattgtc caacaatggt ctatttttgt aagaattaga taaaatgtat acttgatata  2580
aaatagtcaa aaatgtaact cttagtaaca gtaagcttgg catttagata gaccatgaac  2640
acttcgtcag atactctgtt gggtgtttgg gatagcaatt aaaacaaagt attgatagtt  2700
gtatcagagt ctattaggct gcagcaaagg aagtttattc aaaagtataa actatccaag  2760
attatagacg catgatatac ttcacctatt ttttgtctcc ttaatatgta tatatata  2820
tatatatata tatatacaca tatatgtgtg tgtgtatgtg cgtgtgcatg tttaactttt  2880
aattcagtta aaaactttttt tctatttgtt tttcatctgg atatttgatt ctgcatatcc  2940
tagcccaagt gaaccgagaa gatcgagttg taggactaaa ggatagacat gcagaaatgc  3000
attttaaaaa tctgttagct ggaccagacc gacaatgtaa cataattgcc aaagctttgg  3060
ttcgtgacct gaggttatgt ttggtatgaa aaggtcacat tttatattca gtttttctgaa  3120
gttttggttg cataaccaac ctgtggaagg catgaacacc catgtgcgcc ctaaccaaag  3180
gtttttctga atcatccttc acatgagaat tcctaatggg accaagtaca gtactgtggt  3240
ccaacataaa cacacaagtc aggaagaaac tgaggtcctg gcaggttgca ttctcctgat  3300
actttgggag cattttgcag aggaaagaaa cctgaatctc cgcccccttc ccctcagatg  3360
ggcaaaatgc agctcttcct atatgtatac cctgaatctc cgcccccttc ccctcagatg  3420
ccccctgtca gttccccag ctgctaaata tagctgtctg tggctggctg cgtatgcaac  3480
cgcacacccc attctatctg ccctatctcg gttacagtgt agtcctcccc agggtcatcc  3540
tatgtacaca ctacgtattt ctagccacg aggaggggga atcaaacaga aagagagaca  3600
aacagagata tatcggagtc tggcacgggg cacataaggc agcacattag agaaagccgg  3660
cccctggatc cgtctttcgc gtttatttta agcccagtct tccctgggcc acctttagca  3720
gatcctcgtg cgcccccgcc ccctggccgt gaaactcagc ctctatccag cagcgacgac  3780
aagtaaagta aagttcaggg aagctgctct ttgggatcgc tccaaatcga gttgtgcctg  3840
gagtgatgtt taagccaatg tcagggcaag gcaacagtcc ctggccgtcc tccagcacct  3900
ttgtaatgca tatgagctcg ggagaccagt acttaaagtt ggaggccggg gagcccagga  3960
gctggcggag ggcgttcgtc ctgggactgc acttgctccc gtcgggtcgc ccggcttcac  4020
cggaccccgca ggctccccggg gcagggccgg ggccagagct cgcgtgtcgg cgggacatgc  4080
gctgcgtcgc ctctaacctc gggctgtgct cttttttccag gtggcccgcc ggtttctgag  4140
```

```
ccttctgccc tgcggggaca cggtctgcac cctgcccgcg gccacggacc atgaccatga   4200
ccctccacac caaagcatct gggatggccc tactgcatca gatccaaggg aacgagctgg   4260
agccccctgaa ccgtccgcag ctcaagatcc ccctggagcg gcccctgggc gaggtgtacc   4320
tggacagcag caagcccgcc gtgtacaact accccgaggg cgccgcctac gagttcaacg   4380
ccgcggccgc cgccaacgcg caggtctacg gtcagaccgg cctcccctac ggccccgggt   4440
ctgaggctgc ggcgttcggc tccaacggcc tggggggttt ccccccactc aacagcgtgt   4500
ctccgagccc gctgatgcta ctgcacccgc cgccgcagct gtcgcctttc ctgcagcccc   4560
acggccagca ggtgccctac tacctggaga acgagcccag cggctacacg gtgcgcgagg   4620
ccggcccgcc ggcattctac aggtacccgc gcccgcgccg cccgtcgggg tggccgccgc   4680
gcccggcagg agggagggag ggagggaggg agaagggaga gcctaggag ctgcgggagc   4740
cgcgggacgc gcgacccgag ggtgcgcgca gggagcccgg ggcgcgcggc ccagcccggg   4800
ggttctgcgt gcagcccgcg ctgcgttcag agtcaagttc tctcgccggg cagctgaaaa   4860
aaacgtactc tccacccact taccgtccgt gcgagaggca gacccgaaag cccgggcttc   4920
ctaacaaaac acacgttgga aaaccagaca aagcagcagt tatttgtggg ggaaaacacc   4980
tccaggcaaa taaacacggg gcgctttgag tcacttggga aggtctcgct cttggcattt   5040
aaagttgggg gtgtttggag ttagcagagc tcagcagagt tttatttatc cttttaatgt   5100
ttttgtttaa tgtgctcccc aaatttcctt tcatctagac tatttgattg gaaatatgtc   5160
agctatgatg atgactttct gggaagcgat tcctgtcacc cgctttcccc tcctccccac   5220
cccacgtcct ggggctttag agagcgattg ggagttgaat gggtctgatt tcggagttag   5280
ctggctgagt ccgcgctgga gcggattgct ggcatgtgac ttctgacagc cggaaatttg   5340
taggtgtccc gcgagtttaa aacaagccat atggaagcac aagtgcttaa aaataatctc   5400
ctgccagccc agtgacaagc ctgtcccacc cggggagaat gccccggagt ggcgtgcggg   5460
tcagccaggg tctgcgcctc gcagccactg tggaaggagc gcggccggtc caggacacag   5520
gagaccactt tgtgacttca atggcgaagg ttgtgtgtcc tcattttaat ttttttccct   5580
acaagaattg ttctttctcc ctctcctctc cctcccattt tctcttgccc agtttctcct   5640
tttgtttttt gttttttgtt ttcctgatgg gcctgcagag ggattaggtg ggcgcttctg   5700
gtgaacacct tcctaggtgg ccacaggaca ggtgtacccc ggctgtacact ggtatccgga   5760
agggcgccac atggctgggt cctgaattag gcatttccca actgtacact ggtatccgga   5820
ctggtgtccc tatatctttc tgccttgtaa gccgtggacc agttttgtt cagtattctg   5880
tttccaggga tatttatagc agaaggaagg ggactaaagt gcagtttggc cccagaggat   5940
actgaagggc agattctggg ggtattcagt gtgcatcttc agccgccttg gagaaattta   6000
gagcatccca cagccacgca gatccaagct gtctttactc aaaagacaaa caatgaacaa   6060
aacttttaaa ggttggcata tttcaaatta attttacttg ttttaattta gggttaaaac   6120
agagaaaaag gatttcttct gcccaccttt ttttttttaa atggaagaac aaagtacagc   6180
gattaagtct aattccacac aacatttaaa actgcttgat gtgaaggaag gcactggtat   6240
gatgtgaatt ccataacctt atgatggact ccagaaacca ttttcttccc tatttaattt   6300
tcagttcttt tattgcaaat taatgctgct gaatttcaat gggcactaat gagactgctc   6360
cttggtagat tatttactgc cttgctaata attacaaagt gaacctggtc aaatacagag   6420
gggatcgcat cttattcaaa attgttcatc atcccagtga taagtggtat cagtgtaata   6480
tgccctatct tacactttct gcattacatg atattcaaac actcttagaa taataaaaaa   6540
agagacaagg aacttaaaaa ttaaaaaaaa aacttgcaca aatgggactc tgtgtgggaaa   6600
ttcagtttta gaatgatttt tcctgtgttt tatttcccgg attatctttc ctcttttgtt   6660
agaattctgc ctgttattat ccagcaagga aaagaagcat ctatgcaagt tcttcatatg   6720
gacagatatt atttagtatt tttcccctct cagttttct gcttaaatga ctctgggtat   6780
aaaggaaagg attgattggg ctcttttagg aaactttaag tttcttaagt agttctcaaa   6840
agttttgggg ctgaaagcag tgttttcaaa ctgcttgtca tgacccagag ggtcatgaac   6900
tcagtttagt gagtctagaa tatttttaa aaggactaaa atggaaagga atataataga   6960
aaatatcaga gtgcatggta tttcgtaagg ataagttttg t                       7001
```

<210> 5
<211> 11001
<212> DNA
<213> Homo Sapiens

<400> 5

```
ccaagtcaga tgttccccaa ttacctgtgg acaggtcagg catattctga gtctaatttc     60
actccacagg cctaatacct tggagccaga aagcttccag gtaaaaagtc tgaagggggc    120
ctcctcatgt cattagatgg actcctgcat ctccagaaga ttttccacac caggaaagat    180
caaagcacca aggcaattct tcctggcttc ttgggacaac cctaggcttt ggcatgagtg    240
gtctggaagc ctttgcttta gttacaatgc ctatacactc ctggaactgt tttgcagggc    300
ttgtcttcca gcacaattcc tcctccaagc cttactgtag ctacagccca tcagtcctgt    360
ctagtgacaa ccaagaaact aagaactatg tactcacgtt cacctcccca gagtcatttt    420
```

```
ccttcaggac aaagctcagg gccttgtcac tgggccctgc caggagccgc agccgcaggg    480
gctgctcatc atccaacagc ttccgcaagt acactgtgaa ggggaagtaa tgatcagaga    540
cagggccagc tgctcagccc ctgcatgctc aggtgcatgc gtatataccc tcacataggg    600
cagggtgggg tgggaagccc accttggccg tgacgctcag cgcgctcaaa gagtgcaaac    660
ttgcgggggt catccaccac caagaacttt cgcagcaggg cctcaatgac ttcacgtgcc    720
cttgtgcgtg acagcacatg caggtgcttg acagcatcct tgggcaggta aaaggaagtg    780
cggcgcctga cacttgtgcc ccgtcctggg ccccgccggg catcctgcaa ggagggtggc    840
ttcttgctgg agggcacaga gacagggcgc accagcttca gctgaacctt gatgaagcct    900
gtgtaagaac cgtccttgtt ctaaagaaat agagaaacca aaccttgata ataggttcca    960
ggtgagatgt cagtctactt ggggctaggc tgggtatgca caaattactg cttgcgccca   1020
cccaagataa cctcagttgt gaccctctga gtatcaggca catagctggg tacctgctcc   1080
tccccacgcc cccttcctga gcagtcaact caccaagctc atgaagaggt tgctgttgat   1140
ctgggcattg tactccttga tcttctgctc aatctcagct tgagaaaggt caggtgtctc   1200
ccactccaca ggctcgtcct gcaagatggg ccagcatgga cacagggccc ttgaggaacc   1260
cagggcttct ctgaaaaatg gcctctgggg cagtctttgg aaactgactg cctttggccc   1320
cctgtccctg atgtacatat acatagctgg tgcccaccct gaacccacca ctgctcctgg   1380
ttttgcatgc tctgggtgga taagggaaag acagaatcat ttggcttctc tctgctgcct   1440
gcctagggcc tcagcactga atgtagcctt aaggatacca cagaagcagg ggcaactgaa   1500
ggcacatggc caggggccag gaacagctga gggactctga agagggactc tcatttaaag   1560
taaaatcagg ctgggtgtgg tggctcacat ctgcaatccc agcattttgg gaggctaagg   1620
taggaggatc acttgatcct caggagtttg agaccagctt gggcaacata gcaagacctc   1680
atctctacta aaaaagaaa aaaaaaatt agccaggtgt ggtggtgtgc ctgtagtccc   1740
aactgttcag gaggctgagg tgggaggatc gtttgagccc gggagagattgc agctacagta   1800
agctattatc gtgtcactgc actccagcct ggggaactga gtgagaccct gcttcaaaac   1860
acaaaaaaca aaaacaggct gggcacgttg gctcacgcct gtaattctag cactttggga   1920
ggccgaggcg ggtggatcac ctgaggtcag gagtttgaga ccagcctgac caacatggag   1980
aaaccccgtc tctactaaaa atacaaaatt agccaggcgt ggtggcacat gcctgtaatc   2040
ccagctactt aggaggctga ggcaggagaa ttgctcaaac tcgggaggtg gaggttgcag   2100
tgaactgaga tcgtgccatc gcactccagc ctgggcaaca agagcgaaac tcggtctcaa   2160
aaaaaaaaa aatcagtaaa atcacacctc aattgcacat tctgatcaca gcaccctagt   2220
tgagttggag tgagggtttg tcctggagaa ggcagcccat ttttctcctc tgccccggca   2280
cggggccatg acccactgca gggtgagagg agtggagagt ggtgcacatc agtagtccag   2340
ccaccagtgg acagagtagt acttggagcc agttctccat gtctcacaca tagtgagaaa   2400
aatcactgtg acatgatgtt taaccttgac ccaagctgca taaaaggcag ctttaggcca   2460
ggctccaatc tgccagaggt acacaggcag cttcctggtg ggtttctgca cctgcctgtg   2520
ctgtctggag atttggccca aagattttt tttttttga gacgaagcct cactctgtcg   2580
cccaggctgt agtgcagtgg ctggatcttg gctcactgca agttctgcct cctgggttca   2640
agcgattctc ctgcctcaga ctcccgagta gctgggacta caggcgcgtg ccacaacaac   2700
acccggctaa ttttttgtatt tttagtagag atgggatttc accacattgg ccaggttggt   2760
cttgaactcc tgacctcaag tgatccgcct gccttcacct cccaaagtgc tgggattaca   2820
ggcgtgaacc atcgtacccg acccagagat ttttaactcg accactcact ccccacctca   2880
tctagggact ggattcttgc cggaagggtg gagtgtggga cagggcagcc agggctctga   2940
accgactttc ttctcccaga ctcccttggc cccactgcat cagccttact tcctgttgac   3000
gtcagatagg ccctagttag aatgcgagtg tcacagacac agctaagctc agcgctgacc   3060
aatactttgt cccagaagaa ttcccacaag gtttcctgta gaatgatctt gtgcctagcc   3120
caggagagcc agggttctcc ctgactccgc cctggagtcc ccttaagcac ttaaaccatc   3180
tgatggggac aaatggagag gacagatgag ggagcagggt ggagcgtttt agcagaatgc   3240
tccttaccca gaacccgctg ctattctgca gccagcaagg atgtggggct aagaactaag   3300
gccagggcct tacaggaaaa aggtaaaggg ggaggggtgg gaatttaagc tcattttctt   3360
ccccaagtat ccaaaggtct cctggatgga gaagagcact ggagtaaaaa ccccagtaca   3420
aaccttactg gggacagtgg gcaaccttgt cgggttagta aaaacaaatg gtgtgggccc   3480
tggaaaatga gggctggagg ctgtgaataa agcagtggat gtgtttgttc agtacaccaa   3540
cgggaagaag tacccagatg ggaggagtac taggggcagg agaaatgcca gacagactct   3600
agtgccaggg caagaaggaa gatcattttg tttgcagaac agggagggca cagggatggt   3660
gctaacttgt tcttgtgatg gctctgagct cctacctaac aatgagaaag cttgctcctt   3720
cttcccttcc tggatgaccc aggagccctg ggctgggatg cagtgacctc atttccagcc   3780
ccttcccttc tggtgatgaa cctccctatc ttcactcaga aaacagactt ggattagagg   3840
cactgcacag cccttccagg attctaaagg aggaagagtt tctttttctg tttccaaagc   3900
tgcctgctgg aagaggattt caacagccat cccagtctacc actgtccata aaaactgatg   3960
tttcccttct gcaccatagg gacccaccct ccactccagc tggagtgcag tggtgcgatc   4020
gttttttttt tgagacagag tctcgctctg ttttccaggc tggagtgcag tggtgcgatc   4080
ttggctcatt gcaatctctg cctcctgggt tcaagcaatt ctctgcttca gcctcccaag   4140
tagctgggat tacaggtgcc tgccaccaca actggctaat ttttgtatt tttagtcgag   4200
```

```
acggggtttc accattttgg ccaggctggt cttgaactcc tgacctcatg atccacccac    4260
ctcggctttc caaagtgctg ggattaaagg tgtgagccac tgcacctggc ctaaaactga    4320
tgttttttc ttttttttta acatataact tgggacttct cagcctccta ttctttcttt    4380
ttttttttt tttttttga dacagagtct tgctctctca tccaggctgg aatgcagtgg    4440
cccagtctcg actcactgca acctctgtct tctgggttca agtgatactc ctgcctcagc    4500
ctccccagta gctgggatta caggcacaca ccaccatggc cagataattt ttttgtattt    4560
tcagtacaga cggggtttg ctatgttggc ctggcaggtc tcgaactctt ggcctcaagt    4620
gatctgcctg ccttggcctc ccaaaatgct gagattacag gcatgagtca ccaagcccag    4680
ccttcttct tttttttgag acagagcctc accctgtcac ccaggttgga gtgcagtggc    4740
acgatcttgg ctcactgcaa cctttgcctc ccggttgaag tgattcagtc tcccaagtag    4800
ctgggactac agtcacacac caccatgccc ggctaatttt tgtatgttta gtagagatag    4860
ggtttcacca tgttggccag gctgacctcg aattcctgat tgcaaatgat ccacctgcct    4920
tggcctccca aagcattggc attagaggtg tgagccaccg tacttggctt ccttttctat    4980
ttttgagaca gagtctcact ctgtcactca ggctggagtg cagtggcacg atcttggctc    5040
actgcaacct ctgcctccca ggttcaagtg atccttctgc ctcaccctcc caagtagctg    5100
ggattacagg tgtgcacctc cgtggctagc cctccttttc aattggttag tgtcttgtgg    5160
ttttcccacc tttccacagt ggaaaatggc tcaggactga ctgacatgaa gacaagccca    5220
ggggtctaca ctcaactcaa cccttgcacc caagctctgg gctaagattt tggcgtgctg    5280
agcaccaccc attttgtaag gaattttgta aaattttatc tgaagcatca ctcacaactc    5340
cactttcttt acttaaataa ggatttccgc cccatttctg ccaggcatac tgagcttcac    5400
agtccctgtt tcttttcct ggtgcctagg cctggttctc tgagcctggt ggtcacacca    5460
atggcatctg gcacacagtt ctccgataat gggatacct aggaggttcc gagacacctt    5520
acagtcctgg gttagtaacc tggatctctt tttccacctc tttaggcatt ttataatcta    5580
gctttccccc ttcctgtggg taaagtgctc ctgaatgctt atggtccaaa acaagacttc    5640
tttcctatct attcccaaat cttttctccag atccacccta gaggaaggga acagaatctt    5700
ccacattcca gcagctggtg acaggccaga acagggaaga ggtgagggct cagctggctc    5760
catacaggag tgcagatgga ggagcaggat ctctctctgc ctctcaagtt ttcctaaaca    5820
tacttctcaa ttcctggcga ggactcttcc ctctccacat cctcccctag tctccccaag    5880
gagggagcag gagcattcga acgcggaaat cgaggtgcta gtccaaactg ctcggtcggc    5940
tttagtcata gctggataat gcccggctca ggtctaccac aagccataca gctgctttt    6000
ccgtgttcaa cctgtctgtg acagaaacca aggggcccc ggcacccagc atctaggcgg    6060
tggaatcggg gtcttacgca cggttccgcg ggcaggtccc cggccaggac ccgcggggag    6120
ccacgtagcc aggagggtgg ggctgcccac cgacccagga cgcggcaacg gaccggggag    6180
ggcggagctc cagcgaccgc ttcccctccc gcccgccggc acccctggc tcccacctgg    6240
tcccggcgcg gcctgcgagc tagcgaggtt cgcgcggtga agtactgctc gagctccgag    6300
tccgagtcct cttggctgca gtagccactg ctcgtcgtgc tgctccaggt catttcgaaa    6360
gaaggcgcct ccgcctcgcc catagccgta cccgcccgtc ccccagtcct gcgcgtccgt    6420
agccgccaac caccgccccg gtcgcgtgcg tgcgtgtacg cgtgtcagtg tgcgcgtgcg    6480
cccgggccag agccgcgccg caaccgttaa gactgaaacg tagatcgccg ggatctagct    6540
cttgtctcat tggggcagga acgccggggc ggggacacgc acgcttcgcc cccaggaatg    6600
acctcatcgc tccggagctc cactcacaga ccccacctac cacagggaac gggggcgggt    6660
gccagcgtcc gggcaagcgc acaagagtgg cctctggccg gaggcgaggg cgggaaggtg    6720
cgggaagtgc gcgtgcgcgg agcctgggtc agcctgggcc cgggtccgct tgcagcgggt    6780
ggagtacttg cggagccggc aatccaggct cccctcccag cccccgcgca gaattagcct    6840
ctctgtgccg ccgggaaatc ggcaattaga acgctccttg cgcgcggcac ccaggcagcc    6900
ctcgagaatg cctgcactgt ggcctgccca tcctcgccct tcccatacgc cctcggcccc    6960
gcgctcacca cgttcgtgtc ccgctccacc gcgggttccc agcccaggtc ccggggcccg    7020
caacagtcca ggcagacgag cgcgcggcag cggtagtggc aggtgaactt gcaatctgca    7080
gagaggcctg gcggtgaggc ggaggagctc caggtcgggg aaatgtcccg gagattgaag    7140
ggaagcccca gggagagggc cgctgctcgc caggctccgc aggcccgacc tatctcagtg    7200
ggttacctca cactgctacg cggactctaa tgttggccac ctgggcgtct ggaaaccggc    7260
cggaaggcca caggcagaga ggcctgctca acagttggat ctctatcgcc tagcacagaa    7320
cttcccttt cctcattggc aattaaaaaa acaacaacaa aaaactgcgt cttgcttttg    7380
tcacccaggc tggagtgcaa tggcgcgatt tcggctcacc gcaacctccg cctcctgggt    7440
tcaagcgatt cttctgcctc agcctcctga gtacctagga ttacaggcgc cgccaccat    7500
gcccagctaa tttttgtatt tttagtagag acggggtttc accatgttag ccaggctggt    7560
ctcaaactcc tgatctcagg tgatccaccc gcctcggcct tccaaagtgc tgggactaca    7620
ggcttgagcc accgcacccg gcccttcact gggaacgtat atggaataca tctgcccatt    7680
tacttgaagg aaaaactaaa cacctttaac ctacgtctgc cctgtggttg tcacctgtct    7740
ctactcccct cagaccaaga cactggtctc ttcctccagg aaactgcccg ggacagggtc    7800
ccctctaccc actccagcca ggcttgcctc ccagtgttcc attccatt ctcacccct    7860
ctcagcgatc tgtgtactac caaatggaat ccagtgtcc tacactctcc cagtttcct    7920
cagcatcatt tgaagcttgc tcccttgac tcccaggggc tacactctcc cagtttcct    7980
```

```
cctacccccct gcagctcctg ctcagctcct ttgcagattc tgactcaact tccatatctc    8040
acgatgaagt ctgggctcag tcctgatcac tggcctggtc tgtctacatt catctgcccc    8100
agatccacgg ctgaaacact gacctaaacc ctcagactag atcctccgtg ccagtacctt    8160
cactaggatg tctaaaagac gtttcaagtg aacatggcca aaatttaatt cccttttctt    8220
cagcctcact gctacacttg cccagcttcc tctttgcagc aaaaatggcc actaggctcc    8280
cagttactgg agacaaaagc ccaaacttat ctttgatttc tcccttgtct ctacctctga    8340
taaacatgcc caaatcatcc tgcttcttat ctccatggct actttatttc tctttgagaa    8400
cgctgcaatg tcccagcctt gttctttttt ttttttttttt tttttttgag acagagtctc    8460
actctgtcgc caaggctgga gggcagtggc acgatctcgg ctcactgcaa cctccgcctc    8520
ctgtgttcaa gcaattctcc cacctcagcc tcccgagtag ctgggattac aggcacccgc    8580
cactacgcct ggctcatttt tttttatttt ttagtagaca tgaggtttca ccatgttggc    8640
caggctggtc ttgaactcct gacctcaggt gatccacccg cctccgcctt ccgaagtgct    8700
gggattacag gcatgagcca ccgcgctcgg cccttgttc attctttgca ttctgtcaca    8760
actttgtgct cccccccagct gaatttgtga tgtcctcttg taccggatga gagggtctcc    8820
atgcacacac agacctggga cactatccat ccacaagttc ctaaataggc cagagcagtg    8880
atgctcaacc cagactccat gttacaataa tttggggagt ttttaaaatt tactgatgcc    8940
tagggtccac tcccagcagt tgattcaaca ggtctgcggt gggatccagg ctagcgggga    9000
ggactgtaaa agcacccctg gtgattccag ctggtgtcta cccaggggag agcaaccttt    9060
gcttgctggc gattcccagg ggtgcagaag gactgctggg tgtgtggctg cgtgcatatt    9120
ttagcatctg attcactggg tcagaaaagg gtgtttgcta aataaagact caacaaaact    9180
cctgcttgca ggggggcccac caaaggttct aaattttttcc aggctccctc ccataggtgg    9240
taatttccct tcaccctaaa ggttctggag ggggtcatga gtgtttgaga agaggcaagc    9300
ctgggaagat ggactccgag gacagtaggc acaaacccctt tctcaagaag ggccaaggca    9360
ttttaaagat aagaaactta aaatcagcgt attttacat ataagcagcc acctctgctc    9420
atctgtggcc cagatacgag tggagtgcga caagggataa accattttcg cgcactcttc    9480
agcgatgggg cgaaagtaac ggacctagtc ctcgggagct gtccccgccg accccctctg    9540
ccgcgacttg acccgcggcg actgcgctgc cccttggctg ccccttccgc tctcgtaggc    9600
gcgcggggcc actactcacg cgcgcactgc aggcctttgc gcacgacgcc ccagatgaag    9660
tcgccacaga ggtcgcacca cgtgtgcgtg gcgggccccg cgggctggaa gcggtggcca    9720
cggccaggga ccagctgccg tgtggggttg cacgcggtgc cccgcgcgat gcgcagcgcg    9780
ttggcacgct ccagccgggt gcggccctcc ccagcgcgcc cagcgggtgc cagctcccgc    9840
agctcaatga gctcaggctc ccccgacatg gcccggttgg cccgtgctt cgctggcttt    9900
gggcgctagc aagcgcgggc cgggcggggc cacagggcgg gccccgactt cagcgcctcc    9960
cccaggatcc agactgggcg gcgggaagga gctgaggaga gccgcgcaat ggaaacctgg   10020
gtgcagggac tgtggggccc gaaggcgggg ctgggcgcgc tctcgcagag ccccccccgc   10080
cttgcccttc cttccctcct tcgtcccctc ctcacacccc accccggacg gccacaacga   10140
cggcgaccgc aaagcaccac gcggagatac ccgtgtttct ggaggccagc tttactgtgc   10200
tagaggaaga gggtccccac atccggccct ggccctcctg gtccggtttg ctgaagcaac   10260
acacttggcc tacccactgg gtggggcagg aagtctcgag ccttcacttg gggtgaggag   10320
gagggagatc ggtcagcagc tttaccgccc gctctgctct ccactgcgga gactggggct   10380
ccggcagagg ctggaccgtg atcttgaggt tcaggggtgc attctgggtg gattcccttg   10440
gcatgggtgg tcggccctca gcaactgcag ccctcatttg gctctgtcac cctgggctgc   10500
caggacacaa gtctttccat gcttttccca gtgcttgact tggcactccc tgcaggcagg   10560
tgggtattga ggatggcaat gcatgtgggg gatgtgggag tagggcttag aggtccaagg   10620
ttctaggata ccctcacctg gcacagcggg gccgctctgg agccactgcc tctagcacat   10680
aagcctctgc actgcagtaa cacctggggg aaagtcagga aggtctgact ggccctggaa   10740
ccagcctgta ggtctcagcc cacctggggg aaagtcagga aggtctgact ggccctggaa   10800
ggtgggggca ccccacccac atccatgcct cctgcatccc ctccaccctc cctgccattt   10860
ccacaggcct taccttcgcg cctgcagccg caggtcctgc tctgaggggc tgaacacatg   10920
ctggagctgg tgcttggcaa ttgcctgcca cttgcctctg ttttctcgct ccagccgctc   10980
ccagatttct gggatctagg a                                             11001
```

<210> 6
<211> 3501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 6

gcggggttgg tagggggcgtt gttttggtat agttcggggt ttggtagcgg cgggtggggt   60

```
atcggttaag agttgttatc gtcgcgggga ggggagttcg gttcgtcggg atcgtaggta    120
acgggtcgcg gggtttcgcg ggttaggagg ggaacggggt cgggcgggcg agtagcgggt    180
aggggagttt agggttcggt ttcgggtttt gtcgtcggat ttgggggtcg cgaggaagag    240
ttgcgagtcg agggtttggg gtcggcgtat ttttttcgtt ttgtttgtag ttggaaaatt    300
ttttttttaag tttggggcgg cggagtttcg ggggagaagg ggtcggggga gtcgcggagg    360
gaggcgtcgg gttcgcgcgt gtagggttta ggtcgaggtc gggacgcggg tggggcgtag    420
gttcgggtta gggtcgtagt cggttgtgcg tcgtgttcgt tcgggggtt gttttttttt    480
tttttttggga gttgcgtggt tttttttttt ttttatttg ttttttgttt tagtttttg    540
tttcgatata acgttttttt cgcgtcgggt tcggttttcg cgttttgttc gttacggtag    600
tcgttgtttt cgttttttcgc gcggtcgtcg ttcgggtttc gatcgagggt tgatagtttt    660
cggttagggc ggcgttaggg cgggtatcgc gttttttttt ttcgtattat ttttttttaat    720
tggggtaatt ttttcgagg cgggaggcgt tggttttttcg gtttttttttt ttttatttg    780
ggtaaagttt ttcgttttga atgatttttt ttgaagcgga tattttattt aaatcgggta    840
attgtttta aaagggttat tgcgtttgaa tagttttttt ttcggaagtt ttagtattta    900
gttaggtgtt ttggggcgtg taggtcgttt tggtttttttt tttatcggcg gtcgtttatt    960
ttttgttttt tttttttggtt cgggcggggtc ggtttggggtt tttattttag agggtagtcg   1020
gtttttcgtc ggtgtttagg tcgtagggtt gatgttttcg tttagttgag ggaaggggaa   1080
gtggagggga gaagtgtcgg gttggggtta ggcggttagg gcgtcgtacg gtttttattc   1140
ggtcggtgtg tgttttcgta ggagagtgtg ttgggtagac gatgttggat acgatggagg   1200
cgttcggtta ttttaggtag ttgttgttgt agtttaataa ttagcgtatt aagggttttt   1260
tgtgcgacgt gattatcgtg gtgtagaacg ttttttttcg cgcgtataag aacgtgttgg   1320
cggttagtag cgtttattt aagtttttgg tggtgtatga taatttgttt aatttggatt   1380
atgatatggt gagttcggtc gtgtttcgtt tggtgttgga ttttatttat atcggtcgtt   1440
tggttgacgg cgtagaggcg gttgcggtcg cggtcgtggt ttcgggggtt gagtcgagtt   1500
tgggcgtcgt gttggtcgtc gttagttatt tgtagatttt cgatttcgtg gcgttgtgta   1560
agaaacgttt taagcgttac ggtaagtatt gttatttgcg gggcggcggc ggcggcggcg   1620
gcggttacgc gttttatggt cggtcgggtc ggggtttgcg ggtcgttacg tcggttattt   1680
aggtttgtta ttcgtttta gtcgggtttt cgtcgtcgtt tgtcgcggag tcgttttcgg   1740
gtttagaggt cgcggttaat acgtattgcg tcgagttgta cgcgtcggga ttcggttcgg   1800
tcgtcgtatt ttgtgtttcg gagcgtcgtt gttttttttt ttgtggtttg gatttgttta   1860
agaagagttc gtcgggtttc gcggcgttag agcggtcgtt ggttgagcgc gagttgtttt   1920
cgcgttcgga tagttttttt agcgtcggtt tcgtcgttta taaggagtcg ttttttcgttt   1980
tgtcgtcgtt gtcgtcgttg tttttttaga agttggagga ggtcgtatcg tttttcgatt   2040
tatttcgcgg cggtagcggt agttcgggat tcgagttttt cggtcgtttc gacgggttta   2100
gtttttttta tcgttggatg aagtacgagt cgggtttggg tagttatggc gacgagttgg   2160
gtcgggagcg cggtttttttt agcgagcgtt gcgaagagcg tggtggggac gcggtcgttt   2220
cgttcggggg gttttcgttc ggtttggcgt cgtcgtcgcg ttattttggt agtttggacg   2280
ggttcggcgc gggcggcgac ggcgacgatt ataagagtag tagcgaggag atcggtagta   2340
gcgaggattt tagttcgttt ggcggttatt tcgagggtta tttatgttcg tatttggttt   2400
atggcgagtt cgagagtttc ggtgataatt tgtacgtgtg tatttcgtgc ggtaagggtt   2460
tttttagttt tgagtagttg aacgcgtacg tggaggttta cgtggaggag gaggaagcgt   2520
tgtacggtag ggtcgaggcg gtcgaagtgg tcgttggggt cgtcggttta gggttttttt   2580
ttggaggcgg cggggataag gtcgtcgggg tttcgggtgg tttgggagag ttgttgcggt   2640
tttatcgttg cgcgtcgtgc gataagagtt ataaggattc ggttacgttg cggtagtacg   2700
agaagacgta ttggttgatt tatatgcgta gttatttggg ttttaagttt ttcgcgtgcg   2760
agcgtgggat tatgacgcgt tatatcgtt acgcgttagt atcgttttac ggagtatatg cgtatttatt   2820
acgcgtgcgg tatgcggttt acgcgttagt atcgttttac ggagtatatg cgtatttatt   2880
cgggcgagaa gttttacgag tgttaggtgt gcggcggtaa gttcgtatag taacgtaatt   2940
ttattagtta tatgaagatg tacgtcgtgg gtttcgacgg taagggtaag ttcgattttt   3000
cgggtttggg ggggtttttc ggcgttttcg gttcgagta gttgagtttg aagtagtagg   3060
tcgagggcgt ttttgttgtg gttcgtttta cggtcgagta gttgagtttg aagtagtagg   3120
ataaggcggt cgcggtcgag ttgttggcgt agatttacgt cgagttgggt tttagtttcg   3180
tggcgttgga gagtttttat tcgttggtta agtttacggt cgagttgggt tttagtttcg   3240
ataaggcggt cgaggtgttg agttagggcg tttatttggc gttcgttttg tcgttgttgc   3300
ttatcgatcg tttttttttt atttagagcg ttttcgtta gttcgttttg tcgttgttgc   3360
gcggttttgg ttcgtatttt agggagcggc ggggcggcg cgtagggttt attgtgttcg   3420
ggataatcgt agcgtcgtta tagtggcggt tttatttttc ggcggtttta tttggtttta   3480
ttgtttcgtg tttagttcg g                                              3501
```

<210> 7
<211> 3501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 7

```
tcgagttaag gtacgaagta gtgaggttag gtgaggtcgt cgagaggtgg agtcgttatt      60
gtggcgacgt tgcggttgtt tcgggtatag tgggttttgc gcgtcgtttt cgtcgttttt     120
tggggtgcgg gttagggtcg cgtagtagcg atagagcggg ttggcgaggg gcgttttagg     180
tgggagagaa acggtcgatg gttcggtcgt cgggttcggt cgttaggtga gcgttttggt     240
ttagtatttc ggtcgttttg tcggggttga ggtttagttc ggtcgtgaat ttggttagcg     300
ggtagaggtt ttttagcgtt attttggggt cgtgtaggaa gtgcgtggtt tgcgttagta     360
gttcggtcgc ggtcgttttg ttttgttgtt ttaggtttag ttgttcggtc gtgaggcgag     420
ttatagtaaa gacgttttcg gggaagtcga gtttgttttt gtcgtcgggg tcggggacgt     480
cggggagttt ttttaagttc gttagcgttt cggtcgcgtc ggtcgcgttt tttacggcgt     540
gtattttat gtggttgatg aggttgcgtt gttgtgcgaa tttgtcgtcg tatatttggt     600
attcgtaggg ttttcgttc gagtggatgc gtatgtgttt cgtgaggcgg tattggcgcg     660
tgaatcgtat gtcgtacgcg tcgtacgcga agggtttgag gtttaggtgg ttgcgtatgt     720
ggcgcgttat ggttttacgt tgcgtgaatt ttttttcgta gatggtgtat gggtagggtc     780
gggttagtta gtgcgttttt tcgtgttgtc gtagcgtggt cgggtttttg tagtttttgt     840
cgtacgacgc gtagcggtag ggtcgtagta gttttttttag gttattcgga gtttcggcga     900
ttttgttttc gtcgttttta aaggggggtt ttaggtcggc ggttttagcg gttatttcgg     960
tcgtttcggt tttgtcgtat agcgtttttt ttttttttac gtgagttttt acgtgcgcgt    1020
ttagttgttt agagttgggg aagttttttgt cgtacggaat gtatacgtat aggttgttat    1080
cgaagttttc gggttcgtta taggttaggt gcgggtatgg gtagttttcg aggtggtcgt    1140
taggcgggtt ggggttttcg ttgttatcgg tttttcgtt gttgtttttg tagtcgtcgt    1200
cgtcgtcgtt cgcgtcgggt tcgtttaggt tgttagggta gcgcggcggc ggcgttaggt    1260
cgagcggggg ttttttcgggc gagacggtcg cgttttttatt acgttttttcg tagcgttcgt    1320
tggggagtc gcgttttcgg tttagttcgt cgttatagtt atttaggttc ggttcgtgtt    1380
ttatttagcg atagaggaga ttaggttcgt cggggcggtc ggggggttcg ggtttcgggt    1440
tgtcgttgtc gtcgcgaaat gggtcggaag gcggtgcggt ttttttttagt ttttggaagg    1500
gtagcggcgg tagcgacggt agggcgagag gcggtttttt gtaggcggcg gggtcggcgt    1560
tgggagggtt gttcgggcgc ggggtagtt cgcgtttagt tagcggtcgt tttggcgtcg    1620
cggagttcgg cgggtttttt ttggataggt ttaggttata aagaggggag tagcggcgtt    1680
tcgaggtata gagtgcggcg gtcgggtcgg gtttcgacgc gtatagttcg gcgtagtgcg    1740
tgttgatcgc ggttttttggg ttcgagggcg gtttcgcggt aggcggcggc ggaggttcga    1800
ttggggacgg gtagtaggtt tggatgatcg gcgtggcggt tcgtaggttt cggttcggtc    1860
gattataggg cgcgtagtcg tcgtcgtcgt cgtcgtcgtt tcgtaggtgg tagtatttgt    1920
cgtggcgttt gaggcgtttt ttgtatagcg ttacgaggtc ggggatttgt aggtagttgg    1980
cggcggttag tacggcgttt aggttcggtt tagtttttcgg ggttacggtc gcggtcgtag    2040
tcgtttttgc gtcgttagtt aggcggtcgg tgtagatgaa gtttagtatt aggcggaata    2100
cggtcgggtt tattatgtta tggtttaggt tgagtaggtt gttatgtatt attagggatt    2160
tgaggtaggc gttgttggtc gttagtacgt ttttgtgcgc gcggaagagg gcgttttgta    2220
ttacgatgat tacgtcgtat aagaagtttt tggtgcgttg gttgttgagt tgtagtagta    2280
gttgtttgga gtggtcgggc gtttttatcg tgtttagtat cgtttgttta gtatattttt    2340
ttgcggggat atatatcggt cgggtgagag tcgtgcggcg ttttggtcgt ttggtttttag    2400
ttcggtattt ttttttttta tttttttttt ttttagttga gcggggggtat tagtttttgcg    2460
gtttgggtat cggcgaagga tcggttgttt tttggagtgg gagtttaggt cggttcgttc    2520
ggattaggag aaggagtagg aggtgagcgg tcgtcggtgg aggggaggtt agggcggttt    2580
gtacgtttta gggtatttgg ttgggtgttg gggttttcga gaagaaaatt gtttaggcgt    2640
agtgattttt ttggagatag ttattcgatt taagtaaaat gttcgtttta ggaaaagtta    2700
tttagggcgg agaattttat ttaagtaggg agaaagggag tcgaggaatt agcgttttc    2760
gtttcgggag aagttgtttt agttggggga agtgatacgg aggaggggag cgcggtgttc    2820
gtttggcgt cgttttggtc ggggggttgtt aattttcggt cggggttcgg gcggcggtcg    2880
cgcggggagc ggaggtagcg gttgtcgtgg cgggtagagc gcgaaggtcg ggttcggcgc    2940
ggggagggcg ttatatcggg gtaggaggtt gaggtaggaa gtaggtgggg gggaggggggg    3000
agttacgtag ttttttagggg aggggagggg tagcgtttcg ggcgggtacg gcgtatagtc    3060
ggttgcggtt ttgattcggg tttgcgtttt attcgcgttt cggtttcggt ttgggttttta    3120
tacgcgcggg ttcggcgttt ttttttcgcgg tttttttcggt ttttttttt tcggaatttc    3180
gtcgttttaa atttggggaa aagtttttta attgtagata gggcgggagg agtgcgtcgg    3240
ttttaggttt tcggttcgta gtttttttc gcggtttta aattcggcgg tagagttcgg    3300
agtcgagttt tgagttttttt tgttcgttgt tcgttcgttc gattcgtttt ttttttggt    3360
tcgcggggtt tcgcggttcg ttatttgcgg tttcggcggg tcgggttttt ttttcgcgg    3420
```

```
cggtggtagt ttttagtcga tgttttattc gtcgttgtta ggtttcgagt tgtgttaggg    3480
tagcgttttt gttagtttcg t                                              3501
```

<210> 8
<211> 2501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 8

```
ttttttatagt gtaaatgtgt ttttattatt ttttggagta attttatttta aaatcgtttt    60
tagtataaaa tttaaatatt taaatatgat tttgttggtt ttgttttttgt ggtttttattt    120
tttttttttt taaatttagt tagtgtttgt gttgtttgta atgttttttt tttttttgtag    180
gggtcgttat tttaggtttt ggttttttttt tagaaagttt tttttttttt tttttagcgg    240
ggataggggtt tgtttatttt gatattatta gtttatttat atatattggt tataagttta    300
ggttgtatcg ttattgaaag tttattattt gattttgagt agtttgagga ttttattaaa    360
atttaggaga tgtttagtaa atgttgattg aattatgatt gtttttaata tataaacgta    420
agattattta ggaatatttg ttaaaatgtt tttgtttttt gagatttttat tttgggaggt    480
aagtagtggg ggtttaggat tttgtatttt gatagttttt tgatgtttgt atgtagaagt    540
gtagggatta ttatattgat aaatttttat tatttttaag ggggatttttt tttttttaggg    600
gttattttttg gaagttttttt aaggataggg gtcgtatgtt gtttttttag gttagtaatt    660
aaatttagaa aacgtttatt gagtgaatga tgaaacgata ggtgaataga tgaacgtaag    720
gtgtcgagtt aattattttt ttatataagt tttagtagtt tttattgtttt ttagtcgtag    780
aaatggtttt tggaaggtaa gtttttttagc gagtggagtt attttttaatt atatttttta    840
ggattttaag ggagtcgcgc gttttgcgtt tattttttta ttagaaatcg gtaagttatt    900
gattttcgtt tcgttttcgt tattttttcgt ttttttttgt ttcgtagtcg gcgtttagcg    960
gtttttgtttg ttcgtgtgtg tgtcgttgta ggtttttattt atgggtttat cgttgaggtt    1020
cgacgggcgg gtggtattgg ttatcggcgc ggggggtaggt gagtatgcga aggttggagg    1080
tcgcgtttttt tgttgaggcg tagttggttg tttttttcgg gtcggtatac gcgcgtagtc    1140
gtagttgagg ttatttcgtt gaggtggtgg ggaggggaat ggttatttttt gaggtatcgt    1200
attttttgag gaggaaagag tcggaaatat ttggtttttt aagtaggtat agttcgtttt    1260
ttttttagtat ttcggtgtgg gttttttaag gttttgtttg agaggagagg ttaggttggg    1320
ttgttgattg taaaattggg tgaaagtttt ttttgatttt tatttgtggg tatcgattgt    1380
tatttttttt gtaattaatt tttttagatt tttgtttagt tttttaaagg attgaaaagt    1440
cgcgaggggc ggggggttgga attcgttttt tgaagcgtag agatgttagt ttttgaaaag    1500
ttattcggtc gtttagtgtt tgtttttttt tgtcgtaaga ttttaagttc gtgagaggat    1560
ttttttttaaa gagggcgttt gataagagtt ttttttcgtt ggagtttgta tgtttagtaa    1620
gttataattt gttttcgaaa tttattggag ttttggtaga ggttgtaagt ttaaatgcgt    1680
ataggggtta ggcgtatgat ggagaaagaa aatgggagta ggatgggtat atttgaggaa    1740
ttggagagta gagaatttcg aagtggatcg gttagtggga aagttgtttg tattttagga    1800
gcggtaaaat ggaaaattgt tatgtgaaat agttttatttt tttaaagtat aaaaaattaa    1860
aataaaattat ttatattaat atagatgttg tgtagtgaga ttttatatta gttttttatt    1920
agtgggtgat ttttgtaatt tttaagtgta gggatttttga tattatgtat ttttgatttt    1980
ttattggtag tattttatat ttggaaaggt tttaatgtat gaattatttg agttatatat    2040
taaacgttat aaattggaat tttgttaatt aatttttatg tattttata tttgtattga    2100
taaagtggtt ttttatgttg tttttttagaa aatgttttta gtgttgatga atagttaagt    2160
atttttatatt tatagttgtt tggttatttt tgtatgggta tgtatttggg tgtagttata    2220
tttttttaaat gttttttagga aaatattttg tttatatttt gttttttattg taaataatgt    2280
atttttataac gtttggtgtt ttaaattttt tttgatagtt tttggataat ttttatgtag    2340
gaggtttagg gattatattt taagacgttt ttgttatcgt taaggagatt ttttttttta    2400
ggggttatat ttgaaaatta tttaaggata gggattgttt tttttgatat tattagtata    2460
tttatatatg gtatgtagta tattttatat tagtatttag t                        2501
```

<210> 9
<211> 2501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 9

```
attgagtatt ggtgtaaaat gtattgtata ttatgtgtaa gtatgttaat ggtgttaaaa      60
gaagtagttt ttatttttga atgattttta gatatagttt ttgaaaggaa aagttttttt     120
agcgatggta aaaacgtttt agaatgtaat ttttggattt tttgtatgaa aattatttaa     180
gagttgttaa aaaagattta aaatattaag cgttgtaaaa tatattattt ataataaaag     240
taaagtgtaa ataaaatgtt tttttaaaaa tatttagaag gtatgattat atttaaatat     300
atgtttatgt agagatattt gtagtataag aagttatttt attaatatag atatgaaagt     360
ttgaaagtat tttttgaaag gtagtataag aagttatttt attaatatag atatgaaagt     420
atataggaat taattgatag aattttagtt tgtaacgttt aatatataat ttaaataatt     480
tatgtattag ggttttttta ggtataaggt attattagtg gagaattaaa ggtgtataat     540
gttaagattt ttgtatttgg aggttataga ggttatttat tggtgagaaa ttaatgtaaa     600
attttattgt atagtattta tgttggtatg aatggtttgt tttaattttt tgtattttaa     660
aaaatggggt tattttatat aataatttt tatttgtcg ttttgaaat ataggtaatt       720
ttttattgg tcggtttatt tcggaatttt ttgtttttta gtttttttaga tgtgtttatt     780
ttatttttat ttttttttt tattatacgt ttgatttttg tgcgtatttg agtttataat     840
ttttgttaag atttttagtgg atttcgagaa tagattgtga tttgttaagt atataaattt     900
taacgggaa gggttttttat tagacgtttt ttttaaagaa ggttttttta cgaatttaaa     960
attttacgat agagggaaat aaatattgaa cgatcgaatg attttttagg agttgatatt    1020
tttgcgtttt aggggcgaa ttttagtttt cgttttttcgc ggttttttag ttttttaaaa    1080
gattaggtaa agatttaaga gagttaattg taggaagagt aataatcgat gtttatagat    1140
aagggttagg gagaatttt atttagtttt gtaattagta gtttagtttg gtttttttttt    1200
ttaggtagga ttttgggaag tttatatcgg ggtgttgggg agaagcgggt tgtatttgtt    1260
tgagagatta ggtgttttcg gtttttttt ttttaaagaga tgcggtgttt taagaataat    1320
tatttttttt tttattattt tagcggggtg attttagttg cggttgcgcg cgtatgtcgg    1380
ttcgaaaaga gtagttagtt gcgttttagt aaggggcgcg gtttttaatt ttcgtatgtt    1440
tatttgtttt cgcgtcggtg attagtatta ttcgttcgtc gaattttagc ggtgagttta    1500
tgaataaggt ttgtaacgat atatatacga ataagtagag tcgttggacg tcgattgcgg    1560
gataggagga ggcggggaat ggcggggggcg ggacgagggt tagtgatttg tcgatttttg    1620
gtaggaagat gagcgtagag cgcgcggttt tttttggaatt ttgggaaatg tagttaagag    1680
tgattttatt cgttggaaga tttgtttttt aggggttatt tttgcggttg gaagtaatgg    1740
gagttgttag gatttgtgta gaagaatagt taattcgata ttttgcgttt atttatttat    1800
ttgtcgtttt attatttatt taataaacgt tttttgggtt tagttgttga tttagagaaa    1860
tagtatgcgg ttttttatttt tgagggggttt ttagagatag tttttgggaa ggaaagttttt  1920
ttttagggat ggtaaagatt tgttagtgta ataatttttg tatttttata tgtaaatatt     1980
aggggattgt taaaatgtag agttttggat ttttattgtt tattttttaa aatagaattt     2040
taagggggtaa aaatattttg ataagtgttt ttaaatgatt ttgcgtttgt atgttgagaa     2100
tagttatagt ttaattaata tttattgagt attttttgag ttttgatagg attttttaagt     2160
tatttagagt tagatggtaa attttttaata acggtgtagt ttagatttgt gattaatgtg     2220
tgtaagtgag ttaatggtgt taaaataaat agatttattt tcgttgggg agaaagagga      2280
aaaatttttt gaaggaggat taggatttaa agtggcgatt tttgtaaaga aagaagggta      2340
ttataggtag tataaatatt agttaggttt ggggagaaag agggtaaagt tataaaagta      2400
aagttagtaa gattatgttt agatgtttga attttgtgtt gaaacggtt ttaagtagga      2460
ttattttaga gagtggtggg aatatattta tattatggaa a                        2501
```

<210> 10
<211> 2470
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 10

```
aaagatgatt aaaagtttaa ttgtttattt gaagagttga ttttttttatt tttgtaataa      60
agggtatttt tagtagtttt tgtttatttt gtttattcgg ttttttttgt ggttgtgtaa     120
ggttataatt tttgtgtttt agtaaatttg tgtatgttta tttttttttt tgttattatt     180
ttttttttta ttttgtttta ttattttgat gtaaaattat ttgttaattt tatttgaaat     240
gagaaatttt aaggtttata ttatttaaat tttgttagat tttttatttttt gttatatggt    300
ttataatgtg ttgggtattt ttagatttgt ttattaaaaa gatgtaaaat aaaataatga     360
ttattttgt ggattttttt tttatttttg agatgttttt tttggttgta ttatttttt        420
```

```
atttttttgtt tattgattag aggaggggtt ttaattatgg gtgaattttta tatttttattg    480
aagaggttat gttatatgta tatttttata atataattta tatttatata gtatttttat      540
ttttagtata tttttttta ttaattttaa taatattatt gtaagttatg ttgaagtaga       600
ttgtaagtgt ttatttataa attgtgaaat gaattaaaat gaaagggtaa agattaaatt      660
atgattaggt ttgaaattaa tatataagat ttaatttttt ttaattaaag attttttgtag    720
gtgattttttg tttgtaggat tttttttttt ttttagatgt tattggattg tattaggttt    780
attgtagatt ttagtcgttg tagaattaat tagatttaag atgagttttt tgattttttt     840
tggtagagtt ttttaattgt tgaattttaa tattgtcgtg attagttagt gttataattt      900
gtttgtttta ttttgtgtaa tggattttat attatagagg tattttttta atgttaagat      960
gtttaagtat tgtttaagtg taaattattt aatatttttt agttattaag taattaagat     1020
aggtaggatt ttatttgttt taaaatgatt tgatttaaat taaaaagaga atgtggattt     1080
tttgaatttt atttggttaa ttttaatata atttttagta ttttataatt ttttttaaag    1140
ttttttttatt tggttatttt ttgtatttttt tttgttttttt tttttttttt ttagttataa    1200
taattgttag attttgtttt attttttttt gatagttttt attttttaagg ttatttattt    1260
tttttaggta tttttttggtt ttagtttgag tatagtagat tttaagatta tatatgttat    1320
agtataggtt attatagtta attttttgaa taaatgtgat tgaatttttat gttagtaatt    1380
tttatttatt atttttttat taaaaaggtt taaagttttt atttaatgtt tttttttatg    1440
tttatttttgt taaatgattg tttttttaatg atattttaga attttagaat tattttatta    1500
tggaggatgt gtaagattag tttttttatta aataaaaagt gtgaaatgga atatgtaatt    1560
ttattaattt attttggttt taaaattttg tgattattag ataaaattta gaaataaaat    1620
agtattatta atataaataa atttttatta taattatatt ttttaagttt tgtttgtaag    1680
aatgggtaaa atatttttaa aattttgaag aaattattat ttgatagaaa gtttaattta    1740
tttgtgagaa ggtaaatgta tttagatata attaaagttt ttttttttat tttaattta     1800
tttattttga attaagattt tattgtttta ttttttaga tgttgttatt tgaataatat     1860
tgtttttgaga ttaaaaatta gtatattaat ataatttttt ttaaacgttt taagagtttt    1920
gttttttttta tttttttttt taaaaataag tagttattaa attttttagt agtgaatttt    1980
aaaatttttt ttaattttat aggtttaagg gtagttaagg atggttgtag ttttatatga    2040
ttagttgtta aagtaagttg aggtattgaa gatggagaat ttaaattttc gataagagtt    2100
agaagataat tttaattatt ttataaaatt ggaaattgag gtatttaata tgaaggtatt    2160
aagattgtga ttttttaattg tagtttattt atttttattt agtatttttt tttgtaaatt    2220
tgaggtaaga tatttttattt aaaagtgtat tttaaattaa gtaataatat gtaaattttt    2280
ttttgtaaaa gttagtatt atatttttaa ataagatata ttgaatttat ttagtgaatt     2340
atataaagaa aataagtgta aaattttaat ggttagttag ttttttagttt tttttaagat    2400
taaagagaag agattaaata tagtattatt gtattgaggt aaggtttttt gtgtagttta    2460
tagaaattag                                                          2470
```

<210> 11
<211> 2470
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 11

```
ttagttttta tgaattatat agaaaatttt gttttagtat agtgatgtta tatttggttt       60
ttttttttta attttaaaaa gaattaagaa ttaattagtt attggagttt tatatttatt      120
tttttttatat gatttattga atgaatttaa tatatttttat ttaaaaatat aaatgttaat    180
ttttgtaaga aagagtttat atattattgt ttaatttaaa atatattttt aagtaaagtg     240
ttttattttta agtttataag agggaatatt gaataaaaat ggataaatta taattaaaag    300
ttatagtttt gatattttta tattagatgt tttagttttt agttttgtaa gatgattgga     360
attattttttt agttttttgtc gaagatttga gttttttatt tttagtgttt taatttgttt    420
taataattga ttatatgaag ttgtagttat ttttggttat ttttggattt ataaggttaa    480
aaaggatttt gaaatttatt attaaaaaat ttagtggttg tttgttttta aagaaagggg     540
taaaggaaat aaaattttta agacgtttaa gaagaattgt gttaatatgt tagtttttgg     600
ttttaaaata atattgttta agtagtagta tttaagagga tgaaatagtg gagtttttagt    660
ttaagataaa tgaaattaaa atagaagaga gaattttagt tgtgtttgaa tatatttgtt     720
tttttataga tggattaaat tttttattaa gtaataattt ttttaaggtt ttaaagatat    780
tttatttatt tttatäggta aaatttagga aatataatta tgataaaaat ttatttatat    840
tagtaatatt attttatttt tgaatttatat ttgatagtta tagaattttat gagttagaat   900
ggattaatga gattatatat tttattttat atttttttatt tgataaaagg ttaattttat    960
atattttttta tggtgaaata gttttgaagt tttaagatgt tattaaaagg taattatttta   1020
```

```
ataaaatgga tatgaaggag agtattaaat gaagatttta agttttttttg ataggaagat    1080
ggtaaataag aattattaat ataaagttta attatattta tttaaaaggt tgattataat    1140
agtttatgtt atggtatatg tggttttggg atttgttgtg tttaaattga ggttaaaaga    1200
tatttaaaga gaatggatga ttttaggagt agagattgtt aaagagaaat gaagtagagt    1260
ttggtagtta ttatgattgg gaaagaagag gagagataaa gaagatataa aagatagtta    1320
ggtaagagga tttttaggaag aattatagaa tgttaggagt tatattaaga ttaattaagt    1380
aagatttagg agatttatat ttttttttttta gtttaggtta aattattttg gaataaataa    1440
aattttgttt attttaatta tttaatagtt aaaaagtatt aagtagtttg tatttaagta    1500
atatttaaat attttgatat taaaaaaatg tttttgtaat atgaaattta ttatataaaa    1560
taaggtagat aggttgtaat attggttagt tacgataata ttggagttta gtaattggaa    1620
gattttatta aaggaaatta ggggatttat tttagattta gttagtttta taacggttag    1680
aatttatagt aaatttggta taatttaatg atatttgagg aggaagggga gttttgtagg    1740
tagggattat ttataaaagt ttttggttga aaaaaattga gtttttgtgtg ttaatttttag    1800
gtttggttat gatttaattt ttgttttttt attttaattt attttataat ttgtaaatga    1860
atatttataa tttgtttttaa tataatttat agtgatatta ttaggattaa taaaaaaagg    1920
tatgttaaaa ataaaagtat tatgtaaatg taagttatat tatgaaaata tatatgtaat    1980
ataatttttt tagtaagata tagggtttat ttatagttaa gattttttttt ttgattaatg    2040
ggtaaggggt gaagaagtaa tgtagttaaa ggagatattt taaaaataaa ggaaaaattt    2100
ataggagtga ttattatttt gttttatatt ttttttaataa gtaggtttga aaatatttag    2160
tatattataa attatatgat agaggtaggg atttgataga atttgaataa tgtgaatttt    2220
aaaatttttt atttttaaata aaattaatag gtaattttat attaaaataa taaaataaaa    2280
taagagaaaa ggtagtaata gagaaaaaaa tgggtatgta taagtttatt gagatataga    2340
agttataatt ttatataatt ataaaaagag tcggatgggt aagatgagta gagattgtta    2400
aaagtatttt ttattatagg aataaaaaaa ttaattttttt agatgaataa ttaaatttttt    2460
aattattttt                                                            2470
```

<210> 12
<211> 7001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 12

```
aatgtaatgg aaaaagagag attgtaaagt tagaaggttt aggaattgtt ttttgattag      60
gtgtggaagg taagggaaaa ttagttttcg aagaagatag tgagatttta atttgggtgg     120
ttggagagat agtgatgttg ggtatagata cggggaagtt gagaggaata ttatgtttga     180
gaatggtgat ttatatttga ataagtttgt aatgtttagt agatcgttgg aaaagtgggg     240
ttggagatat atttaacgga ggagttagat taatttttat ttttttttat ttgagagagt     300
tagtaagtta cggttggaac gtgtgtgttt agtaggagag ggtagggagg gaagttaaga     360
gagttgggag ttcgagtgaa gttttttgtta aaggtagaag aggaaagtcg gcgtagtata     420
gtatattttt ttatttatgt ttattaagtt tagggataag gtttattaag atgagtttgg     480
aagagaatgt tggagagaaa gtggttaaga aaattgtttt tattgaattt tttgggttaa     540
ttttgattgt aagttttttga ataattaaag tttgtgagga gatagttaat tttttttattt     600
tttttatgtt aatagtgaat aattgtagat tttttttttt tttttttttt tttttttttgt     660
tttttttttt tttttttttg aatattttttg tttttttttg ggattggttt agagtatggg     720
tggttattgt tgatttatag gaggtattat tgttattaat aaagggtaat agttttttttt     780
tttaatattt atttatattt agtatttatt tttaatattg attatggaga gagtttttttt     840
gtgtttaaat attgtaatat tggggggtttt ttaaagtata aaaatatata tttgtatgat     900
ggtattatta atatttttat ggttttttat ttttttttttg tattggtttt aagagttatt     960
tataaatttt ttagtaattg tatagtgttt tagggttaga gatcggttat ttttggtatt    1020
gtgattagag ttatttaata tttaaggtgg tgattaatgt ttggtaataa agttttttatt    1080
gggtgttatg tgttttggga ttttgagcgt gggtatttta ggagtatttt agtattgcgt    1140
gttagtatta tggtcgagag aatagttgag aaagtggtta agaggtggat ttatgtgaac    1200
gttattggga aatgagagat ttcgttttta attacggtta gtgtaattcg aaagtttaaa    1260
attagtttaa aataaaggta tttattttta ttttatgttt atattttagg ttttttaataa    1320
tacgtatttt ttatatgttt atagaaagta gttaattgag ttatttatgg aaaggtttgt    1380
gggtttggtt aacgaagtgg aggagtatta tattttagtt ggaaatatat ttttagaatg    1440
ttaaaatatt tattttaaag tttggttttt tggtgtaatc ggaggtatgg taatgttttt    1500
gtttagagat tggggggttag ggttagtaag gtatttgatt tatatgtatt ttagaaggtt    1560
tttattgtta aattatattt tttcggaaaa attatttatg tttattttg taaatttgat.    1620
```

```
atttatatat ttttgattgg tattttattt tagtcgtaag attatgattt atagtaagtt  1680
tgtttttttt tttgtttggg gtggtagtag aaagtatagg gtattttta gttttaagg   1740
gtaggggtaa aggggttggg gttttttttt tttagtatag ttttttttgg ttgtgttata  1800
ttgtttttg tgagtagata gtaagttttt ttttatttt tattgttatt tatttagcgt   1860
tgtgtagtag tttagttgcg tgtttgtcgg gaggggttgt taagtgtttt gtttattggt  1920
tgttttcga atttttgtta ttttacgtat aaatatattt atatatttt tttgtttagt   1980
ttatatattg agttattcgt atatgcgagt atatttttt tttttttttt attttttcgg   2040
tttttgattt ttataagttt atggaatatt tttggaaaga cgttttgat ttagtagggt   2100
aggtttgttt tgatttttt ttttgtagtt ttagtatttt gagaaagtaa tttatttttt   2160
tggttagtgt ttgtatttta gtagggagat gaggattgtt gtttttatg ggggtatgtg   2220
tgtgtttttt ttttttttta ggatttgtag gattttttgt gttatttgta tataatttgg   2280
taggtttata ttttttaaga gttttatgaa gtgtttttg tatgtgtttt aaaaaggtat   2340
ttgaaaattg aaagtgtgat ttatggaaat taaattattt gtaaaaaatt gttttggaaa  2400
gtaatgattg ttggttataa agggaaatat ttgcgatgta tttaatgtgt ttttaatttt   2460
ttatttgttg ataatttata gttattaatg ttaaattcga ttttggtttt agttatattt   2520
gtatattgtt taataatggt ttatttttgt aagaattaga taaaatgtat atttgatata   2580
aaatagttaa aaatgtaatt tttagtaata gtaagttgg tatttagata gattatgaat   2640
atttcgttag atattttgtt gggtgtttgg gatagtaatt aaaataaagt attgatagtt   2700
gtattagagt ttattaggtt gtagtaaagg aagtttattt aaaagtataa attatttaag   2760
attatagacg tatgatatat tttatttatt ttttgttttt ttaatatgta tatatatata   2820
tatatatata tatatatata tatatgtgtg tgtgtatgtg cgtgtgtatg tttaattttt   2880
aatttagtta aaaatttttt tttatttgtt tttattttgg atatttgatt ttgtatattt   2940
tagtttaagt gaatcgagaa gatcgagttg taggattaaa ggatagatat gtagaaatgt   3000
atttttaaaaa tttgttagtt ggattagatc gataatgtaa tataattgtt aaagttttgg   3060
ttcgtgattt gaggttatgt ttggtatgaa aaggttatat tttatattta gtttttttgaa  3120
gttttggttg tataattaat ttgtggaagg tatgaatatt tatgtgcgtt ttaattaaag   3180
gtttttttga attattttt atatgagaat ttttaatggg attaagtata gtattgtggt   3240
ttaatataaa tatataagtt aggttgagag aattttagaa ggttgtggaa gggtttattt   3300
attttgggag tattttgtag aggaagaaat tgaggttttg gtaggttgta tttttttgat   3360
ggtaaaatgt agtttttttt atatgtatat tttgaatttt cgtttttttt tttttagatg   3420
ttttttgtta gtttttttag ttgttaaata tagttgtttg tggttggttg cgtatgtaat   3480
cgtatatttt attttatttg ttttatttcg gttatagtgt agtttttttt agggttatttt  3540
tatgtatata ttacgtattt ttagttaacg aggaggggga attaaataga aagagagata   3600
aatagagata tatcggagtt tggtacgggg tatataaggt agtatattag agaaagtcgg   3660
ttttttggatt cgttttcgc gtttatttta agtttagttt tttttgggtt attttttagta  3720
gattttcgtg cgttttcgtt ttttggtcgt gaaatttagt ttttatttag tagcgacgat   3780
aagtaaagta aagtttaggg aagttgtttt ttgggatcgt tttaaatcga gttgtgtttg   3840
gagtgatgtt taagttaatg ttagggtaag gtaatagttt ttggtcgttt tttagtatttt  3900
ttgtaatgta tatgagttcg ggagattagt atttaaagtt ggaggttcgg gagtttagga   3960
gttggcggag ggcgttcgtt ttgggattgt atttgttttc gtcgggtcgt tcggtttttat  4020
cggattcgta ggttttcggg gtagggtcgg ggttagagtt cgcgtgtcgg cgggatatgc   4080
gttgcgtcgt tttttaatttc gggttgtgtt tttttttttag gtggttcgtc ggtttttttgag 4140
ttttttttgttt tgcggggata cggtttgtat tttgttcgcg gttacggatt atgattatga   4200
tttttttatat taaagtattt gggatggttt tattgtatta gatttaaggg aacgagttgg   4260
agtttttgaa tcgttcgtag tttaagattt ttttggagcg gtttttgggc gaggtgtatt   4320
tggatagtag taagttcgtc gtgtataatt atttcgaggg cgtcgtttac gagtttaacg   4380
tcgcggtcgt cgttaacgcg taggtttacg gttagatcgg tttttttttac ggtttcgggt   4440
ttgaggttgc ggcgttcggt tttaacggtt tgggggggttt ttttttattt aatagcgtgt    4500
tttcgagttc gttgatgtta ttgtattcgt cgtcgtagtt gtcgtttttt ttgtagtttt   4560
acggttagta ggtgttttat tatttggaga acgagtttag cggttatacg gtgcgcgagg   4620
tcggttcgtc ggtatttttat aggtattcgc gttcgcgtcg ttcgtcgggg tggtcgtcgc   4680
gttcggtagg agggagggag ggagggaggg agaagggaga gtttagggag ttgcgggagt   4740
cgcgggacgc gcgattcgag ggtgcgcgta gggagttcgg ggcgcgcggt ttagttcggg    4800
ggttttgcgt gtagttcgcg ttgcgtttag agttaagttt tttcgtcggg tagttgaaaa   4860
aaacgtattt tttatttatt tatcgttcgt gcgagaggta gattcgaaag ttcgggtttt   4920
ttaataaaat atacgttgga aaattagata aagtagtagt tatttgtggg ggaaaatatt   4980
tttaggtaaa taaatacggg gcgttttgag ttatttggga aggtttcgtt tttggtattt   5040
aaagttgggg gtgtttggag ttagtagagt ttagtagagt tttatttatt tttttaatgt   5100
ttttgtttaa tgtgtttttt aaatttttt ttatttagat tatttgattg gaaatatgtt   5160
agttatgatg atgatttttt agagcgattg ggagttgaat gggtttgatt tcggagttag   5280
ttggttgagt tcgcgttgga gcggattgtt ggtatgtgat ttttgatagt cggaaatttg   5340
taggtgtttc gcgagtttaa aataagttat atggaagtat aagtgtttaa aaataatttt   5400
```

```
ttgttagttt agtgataagt ttgttttatt cggggagaat gtttcggagt ggcgtgcggg      5460
ttagttaggg tttgcgtttc gtagttattg tggaaggagc gcggtcggtt taggatatag      5520
gagattattt tgtgatttta atggcgaagg ttgtgtgttt ttattttaat tttttttttt      5580
ataagaattg tttttttttt tttttttttt ttttttattt tttttttgttt agtttttttt      5640
tttgtttttt gtttttttgtt tttttgatgg gtttgtagag ggattaggtg ggcgttttttg     5700
gtgaatattt ttttaggtgg ttataggata ggtgtatttc ggattgggtt tggaagtttt      5760
agggcgttat atggttgggt tttgaattag gtattttta attgtatatt ggtattcgga       5820
ttggtgtttt tatattttt tgttttgtaa gtcgtggatt agttttttgtt tagtatttttg      5880
tttttaggga tatttatagt agaaggaagg ggattaaagt gtagtttggt tttagaggat       5940
attgaagggt agattttggg ggtatttagt gtgtattttt agtcgttttg gagaaattta       6000
gagtattta tagttacgta gatttaagtt gttttatttt aaaagataaa taatgaataa       6060
aattttaaa ggttggtata ttttaaatta attttatttg ttttaattta gggttaaaat       6120
agagaaaaag gatttttttt gtttatttttt ttttttttaa atggaagaat aaagtatagc      6180
gattaagttt aattttatat aatatttaaa attgtttgat gtgaaggaag gtattggtat      6240
gatgtgaatt ttataatttt atgatggatt ttagaaatta ttttttttttt tatttaatttt    6300
ttagtttttt tattgtaaat taatgttgtt gaattttaat gggtattaat gagattgttt      6360
tttggtagat tatttattgt tttgttaata attataaagt gaatttggtt aaatatagag      6420
gggatcgtat tttatttaaa attgtttatt atttttagtga taagtggtat tagtgtaata      6480
tgtttttattt tatattttttt gtattatata aatttaaat atttttagaa taataaaaaa     6540
agagataagg aatttaaaaa ttaaaaaaaa aattttgtata aatgggattt tgtgtggaaa      6600
tttagtttta gaatgatttt ttttgtgttt tatttttcgg attatttttt ttttttttgtt      6660
agaattttgt ttgttattat ttagtaagga aaagaagtat ttatgtaagt ttttttatatg     6720
gatagatatt atttagtatt ttttttttttt tagtttttttt gtttaaatga ttttgggtat    6780
aaaggaaagg attgattggg ttttttttagg aaattttaag tttttttaagt agtttttaaa    6840
agttttgggg ttgaaagtag tgtttttaaa ttgtttgtta tgatttagag ggttatgaat      6900
ttagtttagt gagtttagaa tatttttttaa aaggattaaa atggaaagga atataataga     6960
aaatattaga gtgtatggta tttcgtaagg ataagttttg t                          7001
```

<210> 13
<211> 7001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 13

```
ataaaattta ttttttacgaa atattatgta ttttgatatt ttttattata ttttttttta       60
ttttagtttt tttaaaaaat attttagatt tattaaattg agtttatgat tttttgggtt        120
atgataagta gtttgaaaat attgttttta gttttaaaat ttttgagaat tatttaagaa        180
atttaaagtt ttttaaaaga gtttaattaa tttttttttt tatatttaga gttatttaag        240
tagaaaaatt gagaggggaa aaatattaaa taatatttgt ttatatgaag aatttgtata        300
gatgtttttt ttttttgttg gataataata ggtagaattt taataaaaga ggaaagataa        360
ttcgggaaat aaaatatagg aaaaattatt ttaaaattga atttttatat agagtttttat       420
ttgtgtaagt ttttttttta attttttaagt tttttgtttt ttttttttatt attttaagag      480
tgtttgaata ttatgtaatg tagaaagtgt aagatagggt atattatatt gatattattt        540
attattggga tgatgaataa ttttgaataa gatgcgattt ttttttgtatt tgattaggtt       600
tattttgtaa ttattagtaa ggtagtaaat aatttattaa ggagtagttt tattagtgtt        660
tattgaaatt tagtagtatt aatttgtaat aaaagaattg aaaattaaat agggaagaaa        720
atggttttttg gagtttatta taaggttatg gaatttatat tatattagtg ttttttttta       780
tattaagtag ttttaaatgt tgtgtggaat tagatttaat cgttgtattt tgtttttttta       840
tttaaaaaaa aaaaggtggg tagaagaaat tttttttttt tgttttaatt ttaaattaaa        900
ataagtaaaa ttaatttgaa atatgttaat ttttaaaagt tttgtttatt gtttgttttt        960
tgagtaaaga tagtttggat ttgcgtggtt gtgggatgtt ttaaattttt ttaaggcggt       1020
tgaagatgta tattgaatat ttttagaatt tgttttttag tatttttttgg ggttaaattg      1080
tattttagtt tttttttttt tgttataaat aggatatta gttcggatat tagtgtatag        1140
tggtttacgg tttataaggt agaaagatat tgtggcgttt tgaagttttt aaatttagtt       1200
ttgggaaatg tttaatttag gatttagtta tgtggcgttt tgaagttttt aaatttagtt       1260
cggggtatat ttgtttttgtg gttatttagg taaaaaataa aaggagaaat tgggtaagag      1380
tttttgtagg tttattagga aaataaaaaa ataattttg tagggaaaaa aattaaaatg        1440
aaaatgggag ggagaggaga gggagaaaga ataatttttg taggaaaaa aattaaaatg        1500
aggatatata attttcgtta ttgaagttat aaagtggttt tttgtgtttt ggatcggtcg       1500
```

```
cgttttttttt atagtggttg cgaggcgtag atttttggttg attcgtacgt tatttcgggg   1560
tatttttttc gggtgggata ggtttgttat tgggttggta ggagattatt tttaagtatt   1620
tgtgttttta tatggtttgt tttaaattcg cgggatattt ataaattttc ggttgttaga   1680
agttatatgt tagtaattcg ttttagcgcg gatttagtta gttaatttcg aaattagatt   1740
tatttaattt ttaatcgttt tttaaagttt taggacgtgg ggtgggagg aggggaaagc   1800
gggtgatagg aatcgttttt tagaaagtta ttattatagt tgatatattt ttaattaaat   1860
agtttagatg aaaggaaatt tggggagtat attaaataaa aatattaaaa ggataaataa   1920
aattttgttg agttttgtta attttaaata ttttaatttt taaatgttaa gagcgagatt   1980
ttttaagtg atttaaagcg tttcgtgttt atttgtttgg aggtgttttt ttttataaat   2040
aattgttgtt ttgtttggtt ttttaacgtt tgttttgtta ggaagttcgg gttttcgggt   2100
ttgttttttcg tacggacggt aagtgggtgg agagtacgtt tttttttagtt gttcggcgag   2160
agaatttgat tttgaacgta gcgcggggttg tacgtagaat tttcgggttg ggtcgcgcgt   2220
ttcgggtttt ttgcgcgtat tttcgggtcg cgcgtttcgc ggttttcgta gtttttttagg   2280
ttttttttttt ttttttttttt ttttttttttt ttttgtcggg cgcggcggtt atttcgacgg   2340
gcggcgcggg cgcgggtatt tgtagaatgt cggcgggtcg gtttcgcgta tcgtgtagtc   2400
gttgggttcg ttttttaggt agtagggtat ttgttggtcg tggggttgta ggaaaggcga   2460
tagttgcggc ggcgggtgta gtagtattag cgggttcgga gatacgttgt tgagtggggg   2520
gaaattttt aggtcgttgg agtcgaacgt cgtagtttta gattcggggt cgtaggggag   2580
gtcggtttga tcgtagattt gcgcgttggc ggcggtcgcg gcgttgaatt cgtaggcggc   2640
gttttcgggg tagttgtata cggcgggttt gttgttgttt aggtatattt cgtttagggg   2700
tcgttttagg gggattttga gttgcggacg gtttaggggt tttagttcgt ttttttggat   2760
ttgatgtagt agggttattt tagatgtttt ggtgtggagg gttatggtta tggttcgtgg   2820
tcgcgggtag ggtgtagatc gtgttttcgt agggtagaag gtttagaaat cggcgggtta   2880
tttggaaaaa gagtatagtt cgaggttaga ggcgacgtag cgtatgtttc gtcgatacgc   2940
gagttttggt ttcggttttg tttcgggagt ttgcgggttc ggtgaagtcg ggcgattcga   3000
cgggagtaag tgtagtttta ggacgaacgt ttttcgttag tttttggggtt ttcgggtttt   3060
taattttaag tattggtttt tcgagtttat atgtattata aaggtgttgg aggacggtta   3120
gggattgttg ttttgttttg atattggttt aaatattatt ttaggtataa ttcgatttgg   3180
agcgatttta aagagtagtt tttttgaatt ttattttatt tgtcgtcgtt gttggataga   3240
ggttgagttt tacggttagg gggcgggggc gtacgaggat ttgttaaagg tggtttaggg   3300
aagattgggt ttaaaataaa cgcgaaagac ggatttaggg gtcggttttt tttaatgtgt   3360
tgttttatgt gtttcgtgtt agatttcgat atatttttgt ttgtttttttt ttttgtttga   3420
ttttttttttt tcgttggtta gaaatacgta gtgtgtatat aggatgattt tggggaggat   3480
tatattgtaa tcgagatagg gtagatagaa tggggtgtgc ggttgtatac gtagttagtt   3540
atagatagtt atatttagta gttgggggaa ttgatagggg gtatttgagg ggaaggggc   3600
ggagatttag ggtatatata taggaagagt tgtattttgt tattaggaga atgtaatttg   3660
ttaggatttt agttttttttt tttgtaaaat gttttaaag tagatagatt ttttttataat   3720
ttttttgagat ttttttttagtt tgatttgtgt gtttatgttg gattatagta ttgtatttgg   3780
ttttattagg aatttttatg tgaaggatga tttagaaaaa tttttggtta gggcgtatat   3840
gggtgtttat gttttttata ggttggttat gtaattaaaa ttttagaaaa ttgaatataa   3900
aatgtgatt ttttatatta aatataattt taggttacga attaaagttt tggtaattat   3960
gttatattgt cggtttggtt tagttaatag atttttaaaa tgtatttttg tatgtttatt   4020
tttttagtttt ataattcgat ttttttcggtt tatttgggtt aggatatgta gaattaaata   4080
tttagatgaa aaataaatag aaaaaagttt ttaattgaat taaaagttaa atatgtatac   4140
gtatatatat atatatatat atgtgtatat atatatatat atatatatat atatatatta   4200
aggagataaa aaataggtga agtatattat gcgtttataa ttttggatag tttatatttt   4260
tgaataaatt ttttttgttg tagtttaata gattttgata taattattaa tattttgttt   4320
taattgttat tttaaatatt taatagagta tttgacgaag tgtttatggt ttatttaaat   4380
gttaagttta ttgttattaa gagttatatt tttgattatt ttatattaag tatatatttt   4440
atttaatttt tataaaaata gattattgtt ggataaatatg taaatgtagt tgaagttaaa   4500
atcgagttta gtattaatga ttatagattg ttagtaaata aagggttaaa aatatattag   4560
gtgtatcgta gatatttttt tttatggtta gtaattatta tttttttaaag taattttttta   4620
tagatgattt aattttttata aattatattt ttaattttta aatgtttttt taaaatatat   4680
gtaaaaagta ttttataggg tttttaaaaa atgtgaattt gttaaattat atgtaaatgg   4740
tataaagaat tttataagtt ttgaaagaaa aaggagatat atatatattt ttatggagaa   4800
tagtaatttt tattttttttg ttaggatata gatattagtt agaaaggtaa gttgtttttt   4860
taaaatgtta aagttataga gagagaaatt aaaataagtt tattttgttg gattaagaac   4920
gttttttttag aaatgtttta tgggtttgta gaagttaagg gtcgagagag tgagaaggaa   4980
ggaaggaatg tgttcgtatg tgcgagtggt ttagtgtgtg aattaggtag agagagtgtg   5040
tggatgtgtt tgtgcgtgga atggtaggga ttcgggaagt agttagtagg tagggtattt   5100
ggtagttttt ttcggtagat acgtagttgg gttattgtat agcgttggat gaatggtagt   5160
ggggagtgag gggagatttg ttgtttgttt atagggagta gtgtggtata gttagagaaa   5220
gttgtattgg ggaggagaaa ttttagtttt tttgtttttta ttttggagg ttggaaagta   5280
```

```
ttttatgttt tttgttgtta ttttaagtaa gaggaaaaat aggtttgttg tgaattatag      5340
ttttacggtt aaaatagaat gttagttaaa agtgtatgga tattaagttt ataaaatagg      5400
atatgggtgg ttttttcgaa agaatataat ttaataataa aagttttttg ggatatatgt      5460
ggattaaatg ttttattggt tttagttttt agttttttgaa tagaggtatt gttatgtttt      5520
cgattgtatt aggaaattag attttggaat aaatgttttg gtattttagg gatgtgtttt      5580
tagttgaaat gtaatatttt tttatttcgt taattaaatt tataaatttt tttatgaata      5640
gtttagttga ttgttttttg taaatatgtg aaaaatacgt attattaaaa gtttaggata      5700
tgaatataag ataaaggtag atattttgt tttaaattga ttttaggttt tcgagttgta        5760
ttgatcgtga ttgggaacga ggttttttat tttttagtgg cgtttatatg gatttatttt      5820
ttgattattt ttttaattat tttttcggtt atagtattaa tacgtaatat tgaggtgttt      5880
ttagagtgtt tacgtttagg gttttaggat atatgatatt taatggaggt tttgttgtta      5940
gatattagtt attatttttgg atattaaatg attttaatta taatgttagg agtggtcggt     6000
ttttggtttt gggatattat gtagttattg agagatttat gagtggtttt tgagattagt       6060
ataaaaaga aatagaaagt tataaaaatg ttaatgatgt tattatgtaa atatatgttt        6120
ttgtgttttg aaagattttt agtattgtag tgtttgagta taggagagtt ttttttatag       6180
ttagtattga aaataaatat tggatataaa taaatattga aaagaaagat tgttattttt       6240
tgttggtgat agtggtgttt tttgtaggtt aataatggtt atttatgttt tagattagtt        6300
ttagaaaaaa gtaagagtat ttagggaggg aggagagagg aataggggaa aggagaagga       6360
aaggaaaggg gatttgtaat tgtttattat tgatatagga agaataagaa ggttagttgt        6420
ttttttatag gttttgattg tttagagatt tataattaaa gttagtttaa gaagtttagt       6480
aaaggtagtt tttttaatta tttttttttt agtatttttt tttaaatttta ttttggtgag      6540
ttttgttttt gggtttggtg agtatgggtg ggaaagtata ttgtgttacg tcgatttttt       6600
tttttgttt ttggtaaaaa ttttattcgg gttttttagtt tttttggttt tttttttat        6660
tttttttgt tggatatata cgttttagtc gtgatttatt ggttttttta ggtgaagaag        6720
ggtaaagatt gatttggttt tttcgttgaa tgtgttttta gttttatttt tttagcggtt        6780
tgttgggtat tgtaggtttg tttaaatatg agttattatt tttaaatatg gtgtttttt         6840
taattttttc gtgtttgtgt ttagtattat tgttttttta gttatttaga ttaaaatttt         6900
attgtttttt tcgagggttg atttttttt gttttttata tttaattaag aggtaatttt         6960
taagtttttt agttttataa tttttttttt tttattgtat t                            7001
```

<210> 14
<211> 11001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 14

```
ttaagttaga tgttttttaa ttatttgtgg ataggttagg tatattttga gtttaatttt        60
attttatagg tttaatattt tggagttaga aagttttttag gtaaaaagtt tgaagggggt       120
tttttatgt tattagatgg attttttgtat ttttagaaga tttttttatat taggaaagat       180
taaagtatta aggtaatttt tttttggtttt ttgggataat tttaggtttt ggtatgagtg      240
gtttggaagt ttttgtttta gttataatgt ttatatattt ttggaattgt tttgtagggt        300
ttgttttta gtataatttt tttttttaagt tttattgtag ttatagttta ttagttttgt       360
ttagtgataa ttaagaaatt aagaattatg tatttacgtt tattttttta gagttatttt       420
tttttaggat aaagtttagg gttttgttat tgggttttgt taggagtcgt agtcgtaggg        480
gttgtttatt atttaatagt tttcgtaagt atattgtgaa ggggaagtaa tgattagaga        540
tagggttagt tgtttagttt ttgtatgttt aggtgtatgc gtatatattt ttatataggg        600
tagggtgggg tgggaagttt attttggtcg tgacgtttag cgcgtttaaa gagtgtaaat        660
ttgcgggggt tatttattat taagaatttt cgtagtaggg ttttaatgat tttacgtgtt        720
tttgtgcgtg atagtatatg taggtgtttg atagtatttt tgggtaggta aaaggaagtg        780
cggcgtttga tatttgtgtt tcgttttggg tttcgtcggg tattttgtaa ggagggtggt       840
tttttgttgg agggtataga gataggcgt attagtttta gttgaatttt gatgaagttt       900
gtgtaagaat cgttttttgtt ttaaagaaat agagaaatta aattttgata ataggtttta     960
ggtgagatgt tagtttattt ggggttaggt gggtatgta taaattattg tttgcgttta       1020
tttaagataa tttttagttgt gatttttttga gtattaggta tatagttggg tatttgtttt    1080
tttttacgtt ttttttttga gtagttaatt tattaagttt atgaagaggt tgttgttgat     1140
ttgggtattg tattttttga tttttttgttt aattttagtt tgagaaaggt taggtgtttt    1200
ttatttttata ggttcgtttt gtaagatggg ttagtatgga tataggggttt ttgaggaatt   1260
tagggttttt ttgaaaaatg gtttttgggg tagttttttgg aaattgattg tttttggttt    1320
tttgttttttg atgtatatat atatagttgg tgtttatttt gaatttatta ttgttttttgg   1380
```

```
ttttgtatgt tttgggtgga taagggaaag atagaattat ttggtttttt tttgttgttt    1440
gtttaggggtt ttagtattga atgtagtttt aaggatatta tagaagtagg ggtaattgaa    1500
ggtatatggt taggggttag gaatagttga gggatttttga agagggattt ttatttaaag    1560
taaaattagg ttgggtgtgg tggtttatat ttgtaatttt agtattttgg gaggttaagg    1620
taggaggatt atttgatttt taggagtttg agattagttt gggtaatata gtaagatttt    1680
attttattta aaaaagaaa aaaaaaaatt agttaggtgt ggtggtgtgt ttgtagtttt    1740
aattgtttag gaggttgagg tgggaggatc gtttgagttc gggagattgt agttatagta    1800
agttattatc gtgttattgt attttagttt ggggaattga gtgagatttt gttttaaaat    1860
ataaaaaata aaaataggtt gggtacgttg gtttacgttt gtaattttag tattttggga    1920
ggtcgaggcg ggtggattat ttgaggttag gagtttgaga ttagtttgat taatatggag    1980
aaatttcgtt tttattaaaa atataaaatt agttaggcgt ggtggtatat gtttgtaatt    2040
ttagttattt aggaggttga ggtaggagaa ttgtttaaat tcgggaggtg gaggttgtag    2100
tgaattgaga tcgtgttatc gtattttagt ttgggtaata agagcgaaat tcggttttaa    2160
aaaaaaaaaa aattagtaaa attatatttt aattgtatat tttgattata gtattttagt    2220
tgagttggag tgagggtttg ttttggagaa ggtagtttat ttttttttttt tgtttcggta    2280
cggggttatg atttattgta gggtgagagg agtggagagt ggtgtatatt agtagtttag    2340
ttattagtgg atagagtagt atttggagtt agtttttttat gttttatata tagtgagaaa    2400
aattattgtg atatgatgtt taattttgat ttaagttgta taaaaggtag tttttaggtta    2460
ggttttaatt tgttagaggt atataggtag tttttttggtg ggtttttgta tttgtttgtg    2520
ttgtttggag atttggttta aagatttttt tttttttttga gacgaagttt tattttgtcg    2580
tttaggttgt agtgtagtgg ttggattttg gtttattgta agtttttgttt tttgggttta    2640
agcgattttt ttgttttaga ttttcgagta gttgggatta taggcgcgtg ttataataat    2700
attcggttaa tttttgtatt tttagtagag atgggatttt attatattgg ttaggttggt    2760
tttgaatttt tgattttaag tgattcgttt gtttttattt tttaaagtgt tgggattata    2820
ggcgtgaatt atcgtattcg atttagagat tttttaattcg attatttatt tttttatttta    2880
tttagggatt ggattttttgt cggaagggtg gagtgtggga tagggtagtt agggttttga    2940
atcgattttt tttttttttaga ttttttttggt tttattgtat tagtttttatt ttttgttgac    3000
gttagatagg ttttagttag aatgcgagtg ttatagatat agttaagttt agcgttgatt    3060
aatatttgt tttagaagaa tttttataag gtttttttgta gaatgatttt gtgtttagtt    3120
taggagagtt agggtttttt ttgatttcgt tttggagttt ttttaagtat ttaaattatt    3180
tgatggggat aaatggagag gatagatgag ggagtagggt ggagcgtttt agtagaatgt    3240
tttttatttta gaattcgttg ttatttttgta gttagtaagg atgtgggggt aagaattaag    3300
gttaggggttt tataggaaaa aggtaaaggg ggaggggtgg gaatttaagt ttatttttttt    3360
ttttaagtat ttaaaggttt tttggatgga gaagagtatt ggagtaaaaa ttttagtata    3420
aattttattg gggatagtgg gtaattttgt cgggttagta aaaataaatg gtgtgggttt    3480
tggaaaatga gggttggagg ttgtgaataa agtagtggat gtgtttgttt agtatattaa    3540
cgggaagaag tatttagatg ggaggagtat taggggtagg agaaatgtta gatagatttt    3600
agtgttaggg taagaaggaa gattattttg tttgtagaat agggaggggta tagggatggt    3660
gttaatttgt ttttgtgatg gttttgagtt tttatttaat aatgagaaag tttgttttttt    3720
ttttttttttt tggatgattt aggagttttg ggttgggatg tagtgatttt atttttagtt    3780
tttttttttt tggtgatgaa tttttttattt tttatttaga aaatagattt ggattagagg    3840
tattgtatag ttttttttagg attttaaagg aggaagagtt tttttttttttg tttttaaagt    3900
tgtttgttgg aagaggattt taatagttat tttagtcgga tgtatagtag gattatggaa    3960
ttttttttttt gtattatagg gatttattttt ttattttatt attgtttata aaaattgatg    4020
gtttttttttt tgagatagag tttcgttttg ttttttaggt tggagtgtag tggtgcgatt    4080
ttggtttatt gtaatttttg tttttttgggt ttaagtaatt ttttgtttta gttttttaag    4140
tagttgggat tataggtgtt tgttattata attggttaat ttttttgtatt tttagtcgag    4200
acggggtttt attatttttgg ttaggttggt tttgaatttt tgattttatg atttatttat    4260
ttcggttttt taaagtgttg ggattaaagg tgtgagttat tgtatttggt ttaaaattga    4320
tgtttttttt tttttttttta atatataatt tgggatttttt tagttttttta tttttttttt    4380
ttttttttttt tttttttttga gatagagttt tgtttgggttttta agtgatattt ttgtttttagt    4440
tttagtttcg atttattgta atttttgttt tttgggttttta agtgatattt ttgtttttagt    4500
tttttttagta gttgggatta taggtatata ttattatggt tagataattt ttttgtatttt    4560
ttagtataga cggggttttg ttatgttggt ttggtaggtt tcgaattttt ggttttaagt    4620
gatttgtttg ttttggtttt ttaaaatgtt gagattatag gtatgagtta ttaagtttag    4680
tttttttttt tttttttgag atagagtttt attttgttat ttaggttgga gtgtagtggt    4740
acgatttttgg tttattgtaa tttttgtttt tcggttgaag tgatttagtt tttttaagtag    4800
ttgggattat agttatatat tattatgttc ggttaatttt tgtatgttta gtagagatag    4860
ggtttttatta tgttggttag gttgatttcg aatttttgat tgtaaatgat ttatttgttt    4920
tggttttttta aagtattggt attagaggt tgagttatcg tatttggttt ttttttttat    4980
ttttgagata gagtttttatt ttgttatttta ggttggagtg tagtggtacg attttggttt    5040
attgtaattt ttgttttttta ggtttaagtg attttttttgt tttatttttt taagtagttg    5100
ggattatagg tgtgtatttt cgtggttagt tttttttttt aattggttag tgttttgtgg    5160
```

```
tttttttatt tttttatagt ggaaaatggt ttaggattga ttgatatgaa gataagttta      5220
ggggtttata tttaatttaa tttttgtatt taagttttgg gttaagattt tggcgtgttg      5280
agtattattt attttgtaag gaattttgta aaattttatt tgaagtatta tttataattt      5340
tattttttttt atttaaataa ggattttcgt tttatttttg ttaggtatat tgagttttat      5400
agtttttgtt tttttttttt ggtgtttagg tttggttttt tgagtttggt ggttatatta      5460
atggtatttg gtatatagtt tttcgataat ggggatattt aggaggtttc gagatatttt      5520
atagttttgg gttagtaatt tggatttttt tttttatttt tttaggtatt ttataatttta      5580
gtttttttttt tttttgtggg taaagtgttt ttgaatgttt atggtttaaa ataagatttt      5640
tttttattt attttaaat tttttttttag atttattttta gaggaaggga atagaatttt      5700
ttatattttta gtagttggtg ataggttaga ataggaaga ggtgagggtt tagttggttt      5760
tatataggag tgtagatgga ggagtaggat ttttttttgt tttttaagtt tttttaaata      5820
tattttttaa tttttggcga ggattttttt tttttatat tttttttttag ttttttttaag      5880
gagggagtag gagtattcga acgcggaaat cgaggtgtta gtttaaattg ttcggtcggt      5940
tttagttata gttggataat gttcggttta ggtttattat aagttatata gttgttttttt      6000
tcgtgtttaa tttgtttgtg atagaaatta aggggggtttc ggtatttagt atttaggcgg      6060
tggaatcggg gtttttacgta cggtttcgcg ggtaggtttt cggttaggat tcgcggggag      6120
ttacgtagtt aggaggggtgg ggttgtttat cgatttagga cgcggtaacg gatcggggag      6180
ggcggagttt tagcgatcgt tttttttttc gttcgtcggt atttttttggt tttttatttgg      6240
tttcggcgcg gtttgcgagt tagcgaggtt cgcgcggtga agtattgttc gagtttcgag      6300
ttcgagtttt tttggttgta gtagttattg ttcgtcgtgt tgttttaggt tatttcgaaa      6360
gaaggcgttt tcgtttcgtt tatagtcgta ttcgttcgtt ttttagtttt gcgcgttcgt      6420
agtcgttaat tatcgtttcg gtcgcgtgcg tgcgtgtacg cgtgttagtg tgcgcgtgcg      6480
ttcgggttag agtcgcgtcg taatcgttaa gattgaaacg tagatcgtcg ggatttagtt      6540
tttgtttttat tggggtagga acgtcggggc ggggatacgg acgttcgtt tttaggaatg      6600
attttatcgt ttcggagttt tatttataga tttttatttat tatagggaac gggggcgggt      6660
gttagcgttc gggtaagcgt ataagagtgg ttttttggtcg gaggcgaggg cgggaaggtg      6720
cgggaagtgc gcgtgcgcgg agtttgggtt agtttgggtt cgggttcgtt tgtagcgggt      6780
ggagtatttg cggagtcggt aatttaggtt ttttttttttag ttttcgcgta gaattagttt      6840
ttttgtgtcg tcgggaaatc ggtaattaga acgttttttg cgcgcggtat ttaggtagtt      6900
ttcgagaatg tttgtattgt ggtttgttta ttttcgtttt ttttatacgt tttcggttttc      6960
gcgtttatta cgttcgtgtt tcgttttatc gcgggttttt agtttaggtt tcggggttcg      7020
taatagttta ggtagacgag cgcgcggtag cggtagtggt aggtgaattt gtaatttgta      7080
gagaggtttg gcggtgaggc ggaggagttt taggtcgggg aaatgtttcg gagattgaag      7140
ggaagtttta gggagagggt cgttgttcgt taggtttcgt aggttcgatt tattttagtg      7200
ggttattttta tattgttacg cggatttttaa tgttggttat ttgggcgttt ggaaatcggt      7260
cggaaggtta taggtagaga ggtttgtttta atagttggat ttttatcgtt tagtatagaa      7320
ttttttttttt ttttattggt aattaaaaaa ataataataa aaaattgcgt tttgttttttg      7380
ttatttaggt tggagtgtaa tggcgcgatt tcggtttatc gtaattttcg tttttttgggt      7440
ttaagcgatt tttttgtttt agtttttttga gtatttagga ttataggcgt tcgttattat      7500
gtttagttaa tttttgtatt tttagtagag acggggtttt attatgttag ttaggttggt      7560
tttaaatttt tgattttagg tgatttattc gtttcggttt tttaaagtgt tgggattata      7620
ggtttgagtt atcgtattcg gttttttttatt gggaacgtat atggaatata tttgttttatt      7680
tatttgaagg aaaaattaaa tatttttaat ttacgtttgt tttgtggttg ttatttgttt      7740
ttattttttt tagattaaga tattggtttt tatatatttt aatttttcgt ttttattttt      7800
tttttattt attttagtta ggtttgtttt tttttttagg aaattgttcg ggatagggtt      7860
tttagcgatt tgtgtattat taaatggaat ttagtgtttt atttttttatt tttattttttt      7920
tagtattatt tgaagtttgt tttttttgat ttttaggggt tatatttttt tagtttttttt      7980
tttattttttt gtagttttttg tttagttttt ttgtagattt tgatttaatt tttatatttt      8040
acgatgaagt ttggggtttag ttttgattat tggtttggtt tgtttatatt tattttgtttt      8100
agatttacgg ttgaaatatt gattttaaatt tttagattag attttttcgtg ttagtatttt      8160
tattaggatg tttaaaagac gttttaagtg aatatggtta aaatttaatt tttttttttt.      8220
tagtttttatt gttatatttg tttagttttt tttttttgtagt aaaaatggtt attaggtttt      8280
tagttattgg agataaaagt ttaaatttat ttttgatttt tttttttgttt ttattttttga      8340
taaatatgtt taaattattt tgttttttat ttttatggtt attttatttt ttttttgagaa      8400
cgttgtaatg tttttagtttt gttttttttt ttttttttttt tttttttttgag atagagtttt      8460
attttgtcgt taaggttgga gggtagtggt acgatttcgg tttattgtaa tttttcgtttt      8520
ttgtgtttaa gtaattttttt tattttagtt tttcgagtag ttgggattat aggtattcgt      8580
tattacgttt ggtttatttt tttttatttt ttagtagata tgaggtttta ttatgttggt      8640
taggttggtt ttgaattttt gattttaggt gatttattcg tttttcgtttt tcgaagtgtt      8700
gggattatag gtatgagtta tcgcgttcgg ttttttgttt atttttttgta ttttgttata      8760
attttgtgtt tttttttagtt gaatttgtga tgtttttttg tatcggatga gaggttttttt      8820
atgtatatat agatttggga tattattttat ttataagttt ttaaataggt tagagtagtg      8880
atgtttaatt tagatttttat gttataataa tttggggagt ttttaaaatt tattgatgtt      8940
```

```
 tagggtttat ttttagtagt tgatttaata ggtttgcggt gggatttagg ttagcgggga     9000
 ggattgtaaa agtattttg gtgattttag ttggtgttta tttaggggag agtaattttt      9060
 gtttgttggc gatttttagg ggtgtagaag gattgttggg tgtgtggttg cgtgtatatt     9120
 ttagtatttg atttattggg ttagaaaagg gtgtttgtta aataaagatt taataaaatt     9180
 tttgtttgta gggggtttat taaaggtttt aaatttttt aggttttttt ttataggtgg       9240
 taattttttt ttattttaaa ggttttggag ggggttatga gtgtttgaga agaggtaagt     9300
 ttgggaagat ggatttcgag gatagtaggt ataaattttt ttttaagaag ggttaaggta     9360
 ttttaaagat aagaaattta aaattagcgt attttatat ataagtagtt attttgtttt       9420
 atttgtggtt tagatacgag tggagtgcga taagggataa attattttcg cgtattttt       9480
 agcgatgggg cgaaagtaac ggatttagtt ttcgggagtt gttttcgtcg attttttttg     9540
 tcgcgatttg attcgcggcg attgcgttgt ttttggttg ttttttcgt tttcgtaggc        9600
 gcgcggggtt attatttacg cgcgtattgt aggttttgc gtacgacgtt ttagatgaag       9660
 tcgttataga ggtcgtatta cgtgtgcgtg gcgggtttcg cgggttggaa gcggtggtta     9720
 cggttaggga ttagttgtcg tgtggggttg tacgcggtgt ttcgcgcgat gcgtagcgcg     9780
 ttggtacgtt ttagtcgggt gcggtttttt ttagcgcgtt tagcgggtgt tagttttcgt     9840
 agtttaatga gtttaggttt tttcgatatg gttcggttgg gttcgtgttt cgttggtttt     9900
 gggcgttagt aagcgcgggt cgggcggggt tatagggcgg gtttcgattt tagcgttttt     9960
 tttaggattt agattgggcg gcgggaagga gttgaggaga gtcgcgtaat ggaaatttgg     10020
 gtgtagggat tgtggggttc gaaggcgggg ttgggcgcgt tttcgtagag tttttttcgt     10080
 tttgtttttt tttttttttt tcgttttttt tttatatttt atttcggacg gttataacga     10140
 cggcgatcgt aaagtattac gcggagatat tcgtgttttt ggaggttagt tttattgtgt     10200
 tagaggaaga gggttttat attcggtttt ggtttttttg gttcggtttg ttgaagtaat      10260
 atatttggtt tatttattgg gtggggtagg aagtttcgag tttttatttg gggtgaggag     10320
 gagggagatc ggttagtagt tttatcgttc gtttgttttt ttattgcgga gattgggtt      10380
 tcggtagagg ttggatcgtg attttgaggt ttaggggtgt attttgggtg gatttttttg     10440
 gtatgggtgg tcggttttta gtaattgtag tttttatttg gttttgttat tttgggttgt     10500
 taggatataa gtttttttat gtttttttta gtgtttgatt tggtattttt tgtaggtagg     10560
 tgggtattga ggatggtaat gtatgtgggg gatgtgggag tagggtttag aggtttaagg     10620
 ttttaggata ttttattttg tagtaatatt atttattttg gtatcgtgag tagcgtttag     10680
 aagttttgt attgtagtaa gtatagcggg gtcgttttgg agttattgtt tttagtatat      10740
 ttagtttgta ggttttagtt tatttgggg aaagttagga aggtttgatt ggttttggaa       10800
 ggtgggggta ttttatttat atttatgttt tttgtatttt ttttattttt tttgttattt     10860
 ttataggttt tattttcgcg tttgtagtcg taggtttgt tttgaggggt tgaatatatg       10920
 ttggagttgg tgtttggtaa ttgtttgtta tttgttttg ttttttcgtt ttagtcgttt      10980
 ttagattttt gggatttagg a                                               11001
```

<210> 15

<211> 11001

<212> DNA

<213> Artificial Sequence

<220>

<223> chemically treated genomic DNA (Homo sapiens)

<400> 15

```
 ttttagattt tagaaatttg ggagcggttg gagcgagaaa atagaggtaa gtggtaggta      60
 attgttaagt attagtttta gtatgtgttt agttttttag agtaggattt gcggttgtag     120
 gcgcgaaggt aaggtttgtg gaaatggtag ggagggtgga ggggatgtag gaggtatgga     180
 tgtgggtggg gtgttttat tttttagggt tagttagatt ttttttgattt ttttttaggt     240
 gggttgagat ttataggttg gatgtgttag aggtagtggt tttagagcgg tttcgttgtg     300
 tttattgtag tgtagaggtt tttaagcgtt gtttacgatg ttagaatgag tggtattgtt     360
 gtaggtgagg gtattttaga attttggatt tttaagtttt attttatat tttttatatg      420
 tattgttatt tttaatattt atttgtttgt agggagtgtt aagttaagta ttgggaaaag     480
 tatggaaaga tttgtgtttt ggtagtttag ggtgatagag ttaaatgagg gttgtagttg     540
 ttgagggtcg attatttatg ttaagggaat ttattagaa tgtattttg aatttttaaga      600
 ttacggttta gttttgtcg gagtttttagt tttcgtagtg gagagtagag cgggcggtaa     660
 agtgttgat cgattttttt tttttttatt ttaagtgaag gttcgagatt ttttgtttta      720
 tttagtgggt aggttaagtg tgttgtttta gtaaatcgga ttaggaggt tagggtcgga      780
 tgtgggatt ttttttttt agtatagtaa agttggtttt tagaaatacg ggtattttcg       840
 cgtggtgttt tgcggtcgtc gtcgttgtgg tcgttcgggg tggggtgtga ggagggacg      900
 aaggagggaa ggaagggtaa ggcggggggg gttttgcgag agcgcgttta gtttcgtttt     960
 cgggttttat agttttgta tttaggtttt tattgcgcgg tttttttttag tttttttcg     1020
```

```
tcgtttagtt  tggattttgg  gggaggcgtt  gaagtcgggg  ttcgttttgt  ggtttcgttc     1080
ggttcgcgtt  tgttagcgtt  taaagttagc  gaagtacggg  tttaatcggg  ttatgtcggg     1140
ggagtttgag  tttattgagt  tgcgggagtt  ggtattcgtt  gggcgcgttg  ggaagggtcg      1200
tattcggttg  gagcgtgtta  acgcgttgcg  tatcgcgcgg  ggtatcgcgt  gtaatttat      1260
acggtagttg  gtttttggtc  gtggttatcg  ttttagttc   gcggggttcg  ttacgtatac     1320
gtggtgcgat  ttttgtggcg  attttatttg  gggcgtcgtg  cgtaaaggtt  tgtagtgcgc     1380
gcgtgagtag  tggtttcgcg  cgtttacgag  agcggaaggg  gtagttaagg  ggtagcgtag     1440
tcgtcgcggg  ttaagtcgcg  gtagaggggg  tcggcgggga  tagtttttcga ggattaggtt     1500
cgttattttc  gttttatcgt  tgaagagtgc  gcgaaaatgg  tttatttttt  gtcgtatttt     1560
attcgtattt  gggttataga  tgagtagagg  tggttgttta  tatgtaaaaa  tacgttgatt     1620
ttaagttttt  tattttttaaa atgttttggt  ttttttttgag aaagggtttg  tgtttattgt     1680
tttcggagtt  tattttttta  ggtttgtttt  tttttaaata  tttatgattt  ttttttagaat    1740
ttttagggtg  aagggaaatt  attatttatg  ggagggagtt  tggaaaaatt  tagaattttt     1800
ggtgggtttt  ttgtaagtag  gagtttttgtt sagtttttat  ttagtaaata  ttttttttttg     1860
atttagtgaa  ttagatgtta  aaatatgtac  gtagttatat  atttagtagt  tttttttgtat    1920
ttttgggaat  cgttagtaag  taaaggttgt  tttttttttgg gtagatatta  gttggaatta     1980
ttaggggtgt  ttttatagtt  tttttcgtta  gtttggattt  tatcgtagat  ttgttgaatt     2040
aattgttggg  agtggatttt  aggtattagt  aaattttaaa  aatttttttaa attattgtaa     2100
tatggagttt  gggttgagta  ttattgtttt  ggtttattta  ggaatttgtg  gatggatagt     2160
gttttaggtt  tgtgtgtgta  tggagatttt  tttattcggt  ataagaggat  attataaatt     2220
tagttggggg  gagtataaag  ttgtgataga  atgtaaagaa  tgaataaggg  gtcgagcgcg     2280
gtggtttatg  tttgtaatttt tagtatttcg  gaaggcggag  gcgggtggat  tatttgaggt     2340
taggagttta  agattagttt  ggttaatatg  gtgaaatttt  atgtttatta  aaaaataaaa     2400
aaaaatgagt  taggcgtagt  ggcgggtgtt  tgtaatttta  gttattcggg  aggttgaggt     2460
gggagaattg  tttgaatata  ggaggcggag  gttgtagtga  gtcgagatcg  tgttattgtt     2520
ttttagtttt  ggcgatagag  tgagatttttg ttttaaaaaa  aaaaaaaaaa  aaaaaaagaa     2580
taaggttggg  atattgtagc  gtttttaaag  agaaataaag  tagttatgga  gataagaagt     2640
aggatgattt  gggtatgttt  attagaggta  gagataaggg  agaaattaaa  gataagtttg     2700
ggttttttgtt tttagtaatt  gggagtttag  tggttatttt  tgttgtaaag  aggaagttgg     2760
gtaagtgtag  tagtgaggtt  gaagaaaagg  gaattaaatt  ttggttatgt  ttatttgaaa     2820
cgttttttag  atattttagt  gaaggtattg  gtacggagga  tttagtttga  gggtttaggt     2880
tagtgtttta  gtcgtggatt  tggggtagat  gaatgtagat  agattaggtt  agtgattagg     2940
attgagttta  gattttatcg  tgagatatgg  aagttgagtt  agaatttgta  aaggagttga     3000
gtaggagttg  taggggggtag gaggaaaatt  gggagagtgt  agttttttggg agttaaaggg     3060
agtaagtttt  aaatgatgtt  gaggggggtga gaatggagaa  tggaatattg  gatttttattt    3120
ggtagtatat  agatcgttga  ggatttttgtt tcgggtagtt  ttttggagga  agaggtaagt     3180
ttggttggag  tgggtagagg  ggagagtgaa  ggcgaaggat  tagagtgtat  agagattagt     3240
gttttggttt  gaggggagta  gagataggtg  ataattatag  ggtagacgta  ggttaaaggt     3300
gtttagtttt  tttttttaagt  aaatgggtag  atgtatttta  tatacgtttt  tagtgaaggg     3360
tcgggtgcgg  tggtttaagt  ttgtagtttt  agtattttgg  aaggtcgagg  cgggtggatt     3420
atttgagatt  aggagtttga  gattagtttg  gttaatatgg  tgaaatttcg  ttttttattaa    3480
aaatataaaa  attagttggg  tatggtggcg  ggcgtttgta  attttaggta  tttaggaggt     3540
tgaggtagaa  gaatcgtttg  aatttaggag  gcggaggttg  cggtgagtcg  aaatcgcgtt     3600
attgtatttt  agtttggggtg ataaaagtaa  gacgtagttt  tttgttgttg  tttttttttaat   3660
tgttaatgag  gaaaggggaa  gttttgtgtt  aggcgataga  gatttaattg  ttgagtaggt     3720
tttttttgttt  gtggtttttttc ggtcggtttt  tagacgttta  ggtggttaat  attagagttc    3780
gcgtagtagt  gtgaggtaat  ttattgagat  aggtcgggtt  tgcggagttt  ggcgagtagc     3840
ggtttttttt  ttggggttttt  ttttttaattt  tcgggatatt  ttttcgattt  ggagttttttt   3900
cgttttatcg  ttaggttttt  ttgtagattg  taagtttatt  tgttattatc  gttgtcgcgc     3960
gttcgtttgt  ttggattgtt  gcgggtttcg  ggatttgggt  tgggaattcg  cggtggagcg     4020
ggatacgaac  gtggtgagcg  cggggtcgag  ggcgtatggg  aaggcgcagg  atgggtaggt     4080
tatagtgtag  gtattttcga  gggttgtttg  ggtgtcgcgc  gtaaggagcg  ttttaattgt     4140
cgattttttcg gcggtataga  gaggttaatt  ttgcgcgggg  gttgggaggg  gagtttggat     4200
tgtcggtttc  gtaagtattt  tattcgttgt  aagcggattc  gggtttaggt  tgatttaggt     4260
ttcgcgtacg  cgtattttttc gtattttttttc gttttcgttt  tcggttagag  gttattttttg   4320
tgcgtttgtt  cggacgttgg  tattcgtttt  cgttttttgt  ggtaggtggg  gtttgtgagt     4380
ggagtttcgg  agcgatgagg  ttattttttgg gggcgaagcg  tgcgtgtttt  cgtttcggcg     4440
ttttttgtttt  aatgagataa  gagttagatt  tcggcgattt  acgttttagt  tttaacggtt     4500
gcggcgcggt  tttggttcgg  gcgtacgcgt  atattgatac  gcgtatacgt  acgtacgcga     4560
tcggggcggt  ggttggcggt  tacggacgcg  taggattggg  ggacgggcgg  gtacggttat     4620
gggcgaggcg  gaggcgtttt  ttttcgaaat  gatttggagt  agtacgacga  gtagtggtta     4680
ttgtagttaa  gaggattcgg  attcggagtt  cgagtagtat  tttatcgcgc  gaatttcgtt     4740
agttcgtagg  tcgcgtcggg  attaggtggg  agttaggggg  tgtcggcggg  cgggaggggga    4800
```

```
agcggtcgtt ggagtttcgt ttttttcggt tcgttgtcgc gttttgggtc ggtgggtagt   4860
tttatttttt tggttacgtg gttttcgcg gggttttggtc ggggatttgt tcgcggaatc   4920
gtgcgtaaga tttcgatttt atcgtttaga tgttgggtgt cggggttttt ttggttttg   4980
ttatagatag gttgaatacg gaaaaagtag ttgtatggtt tgtggtagat ttgagtcggg   5040
tattatttag ttatgattaa agtcgatcga gtagtttgga ttagtatttc gattttcgcg   5100
ttcgaatgtt tttgttttt ttttggggag attaggggag gatgtggaga gggaagagtt   5160
ttcgttagga attgagaagt atgtttagga aaatttgaga ggtagagaga gattttgttt   5220
ttttatttgt atttttgtat ggagttagtt gagttttat tttttttttg ttttggtttg   5280
ttattagttg ttggaatgtg gaagattttg tttttttttt ttagggtgga tttggagaaa   5340
gatttgggaa tagataggaa agaagttttg ttttggatta taagtattta ggagtatttt   5400
atttatagga aggggggaaag ttagattata aaatgtttaa agaggtggaa aaagagattt   5460
aggttattaa tttaggattg taaggtgttt cggaattttt taggtatttt tattatcgga   5520
gaattgtgtg ttagatgtta ttggtgtgat tattaggttt agagaattag gtttaggtat   5580
taggaaaaag aaataggggat tgtgaagttt agtatgtttg gtagaaatgg ggcggaaatt   5640
tttatttaag taaagaaagt ggagttgtga gtgatgtttt agataaaatt ttataaaatt   5700
ttttataaaa tgggtggtgt ttagtacgtt aaaatttttag tttagagttt gggtgtaagg   5760
gttgagttga gtgtagattt ttgggtttgt ttttatgtta gttagttttg agttattttt   5820
tattgtggaa aggtgggaaa attataagat attaattaat tgaaaaggag ggttagttac   5880
ggaggtgtat atttgtaatt ttagttattt gggagggtga ggtagaagga ttatttgaat   5940
ttgggaggta gaggttgtag tgagttaaga tcgtgttatt gtattttagt ttgagtgata   6000
gagtgagatt ttgtttttaaa aatagaaaag gaagttaagt acggtggttt atatttttaa   6060
tgttaatgtt ttgggaggtt aaggtaggtg gattatttgt aattaggaat tcgaggttag   6120
tttggttaat atggtgaaat tttattttta ttaaatatat aaaaattagt cgggtatggt   6180
ggtgtgtgat tgtagtttta gttatttggg agattgaatt attttaatcg ggaggtaaag   6240
gttgtagtga gttaagatcg tgttattgta ttttaatttg ggtgataggg tgaggttttg   6300
ttttaaaaaa aagaaagaag gttgggtttg gtgatttatg tttgtaattt tagtattttg   6360
ggaggttaag gtaggtagat tatttgaggt taagagttcg agatttgtta ggttaatata   6420
gtaaaatttc gtttgtattg aaaatataaa aaaattattt ggttatggtg gtgtgtgttt   6480
gtaattttag ttattgggga ggttgaggta ggagtattat ttgaatttag aagatagagg   6540
ttgtagtgag tcgagattgg gttattgtat tttagtttgg atgagagagt aagattttgt   6600
tttaaaaaaa aaaaaaaaaa aaaagaaaga ataggaggtt gagaagtttt aagttatatg   6660
ttaaaaaaaa agaaaaaaat attagtttta ggttaggtgt agtggtttat atttttaatt   6720
ttagtatttt ggaaagtcga ggtgggtgga ttatgaggtt aggagtttaa gattagtttg   6780
gttaaaatgg tgaaatttcg tttcgattaa aaatataaaa aattagttag ttgtggtggt   6840
aggtatttgt aattttagtt atttgggagg ttgaagtaga gaattgtttg aatttaggag   6900
gtagagattg taatgagtta agatcgtatt attgtatttt agtttggaaa atagagcgag   6960
attttgtttt aaaaaaaaaa ttattagttt ttatggatag tggtagagtg gagggtgggt   7020
ttttatggtg tagaagggaa attttatggt tttgttgtgt attcgattgg gatggttgtt   7080
gaaatttttt tttagtaggt agttttggaa atagaaaaag aaattttttt tttttagaa   7140
ttttggaagg gttgtgtagt gtttttaatt taagtttgtt ttttgagtga agatagggag   7200
gtttattatt agaagggaag gggttggaaa tgaggttatt gtattttagt ttagggtttt   7260
tgggttattt aggaagggaa gaaggagtaa gttttttttat tgttaggtag gagtttagag   7320
ttattataag aataagttag tattatttttt gtgttttttt tgttttgtaa ataaaatgat   7380
tttttttttt gttttggtat tagagtttgt ttggtatttt ttttgttttt agtattttt   7440
ttatttgggt attttttttc gttggtgtat tgaataaata tatttattgt tttatttata   7500
gttttttagtt tttatttttt agggtttata ttatttgttt ttattaattc gataaggttg   7560
tttattgttt ttagtaaggt ttgtattggg gttttttattt tagtgttttt ttttattttag   7620
gagatttttg gatatttggg gaagaaaatg agtttaaatt tttatttttt tttttttatt   7680
ttttttttgt aaggttttgg ttttagtttt tagttttata tttttgttgg ttgtagaata   7740
gtagcgggtt ttgggtaagg agtattttgt taaaacgttt tattttgttt ttttatttgt   7800
tttttttatt tgttttttatt agatggttta agtgtttaag gggatttttag ggcggagtta   7860
gggagaattt tggtttttttt gggttaggta taagattatt ttataggaaa ttttgtggga   7920
atttttttgg gataaagtat tggttagcgt tgagtttagt tgtgtttgtg atattcgtat   7980
tttaattagg gtttatttga cgttaatagg aagtaaggtt gatgtagtgg ggttaaggga   8040
gtttgggaga agaaagtcgg tttagagttt tggttgtttt gttttatatt ttatttttc   8100
ggtaagaatt tagtttttag atgaggtggg gagtgagtgg tcgagttaaa aatttttggg   8160
tcgggtacga tggtttacgt ttgtaatttt agtattttgg gaggtgaagg taggcggatt   8220
atttgaggtt aggagtttaa gattaatttg gttaatgtgg tgaaatttta tttttattaa   8280
aaatataaaa attagtcggg tgttgttgtg gtacgcgttt gtagttttag ttattcggga   8340
gtttgaggta ggagaatcgt ttgaatttag gaggtagaat ttgtagtgag ttaagatttta   8400
gttattgtat tatagtttgg gcgatagagt gaggtttcgt tttaaaaaaa aaaaaaattt   8460
ttgggttaaa tttttagata gtataggtag gtgtagaaat ttattaggaa gttgtttgtg   8520
tatttttggt agattggagt ttggtttaaa gttgtttttt atgtagtttg ggttaaggtt   8580
```

```
aaatattatg ttatagtgat ttttttattt atgtgtgaga tatggagaat tggttttaag      8640
tattattttg tttattggtg gttggattat tgatgtgtat tattttttat tttttttatt      8700
ttgtagtggg ttatggtttc gtgtcggggt agaggagaaa aatgggttgt ttttttttagg     8760
ataaattttt attttaattt aattagggtg ttgtgattag aatgtgtaat tgaggtgtga      8820
ttttattgat tttttttttt tttgagatcg agtttcgttt ttgttgttta ggttggagtg      8880
cgatggtacg attttagttt attgtaattt ttattttttcg agtttgagta attttttttgt    8940
tttagttttt taagtagttg ggattatagg tatgtgttat tacgtttggt taattttgta      9000
tttttagtag agacggggtt tttttatgtt ggttaggttg gttttaaatt tttgattttta     9060
ggtgatttat tcgtttcggt tttttaaagt gttagaatta taggcgtgag ttaacgtgtt      9120
tagtttgttt ttgtttttttg tgttttgaag tagggttttta tttagttttt taggttggag    9180
tgtagtgata cgataatagt ttattgtagt tgtaatttttt cgggtttaaa cgattttttt     9240
attttagttt tttgaatagt tgggattata ggtatattat tatatttggt taatttttttt    9300
ttttttttttt ttagtagaga tgaggttttg ttatgttgtt taagttggtt ttaaattttt    9360
gaggattaag tgattttttt attttagttt tttaaaatgt tgggattgta gatgtgagtt      9420
attatattta gtttgatttt attttaaatg agagtttttt tttagagttt tttagttgtt      9480
tttggttttt ggttatgtgt ttttagttgt ttttgttttt gtggtatttt taaggttata      9540
tttagtgttg aggttttagg taggtagtag agagaagtta aatgattttg ttttttttttt     9600
atttatttag agtatgtaaa attaggagta gtggtgggtt tagggtgggt attagttatg      9660
tatatgtata ttagggatag ggggttaaag gtagttagtt tttaaagatt gttttagagg      9720
ttatttttta gagaagtttt gggttttttta agggttttgt gtttatgttg gtttattttg      9780
taggacgagt ttgtggagtg ggagatattt gatttttttt aagttgagat tgagtagaag      9840
attaaggagt ataatgttta gattaatagt aattttttta tgagtttggt gagttgattg      9900
tttaggaagg gggcgtgggg aggagtaggt atttagttat gtgtttgata tttagagggt      9960
tataattgag gttattttgg gtgggcgtaa gtagtaattt gtgtatattt agtttagttt      10020
taagtagatt gatattttat ttggaattta ttattaaggt ttggtttttt tatttttttta    10080
gaataaggac ggtttttata taggtttttat taaggtttag ttgaagttgg tgcgttttgt     10140
ttttgtgttt tttagtaaga agttattttt tttgtaggat gttcggcggg gtttaggacg      10200
gggtataagt gttaggcgtc gtattttttt ttatttgttt aaggatgttg ttaagtattt      10260
gtatgtgttg ttacgtataa gggtacgtga agttattgag gttttgttgc gaaagttttt      10320
ggtggtggat gattttcgta agtttgtatt ttttgagcgc gttgagcgtt acggttaagg      10380
tgggtttttt attttatttt gttttatgtg agggtatata cgtatgtatt tgagtatgta      10440
ggggttgagt agttggtttt gttttttgatt attatttttt ttttatagtg tatttgcgga     10500
agttgttgga tgatgagtag tttttgcggt tgcggttttt ggtagggttt agtgataagg      10560
ttttgagttt tgtttttgaag gaaaatgatt ttggggaggt gaacgtgagt atatagtttt     10620
tagtttttttg gttgttatta gataggattg atgggttgta gttatagtaa ggtttggagg     10680
aggaattgtg ttggaagata agttttgtaa aatagtttta ggagtgtata ggtattgtaa      10740
ttaaagtaaa ggttttttaga ttatttatgt taaagtttag ggttgtttta agaagttagg     10800
aagaattgtt ttggtgtttt gattttttttt ggtgtggaaa attttttgga gatgtaggag     10860
tttatttaat gatatgagga ggtttttttt agattttttta tttggaagtt ttttggtttt     10920
aaggtattag gtttgtggag tgaaattaga tttagaatat gtttgatttg tttataggta      10980
attggggaat atttgatttg g                                                11001
```

<210> 16
<211> 3501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 16

```
gtggggttgg taggggtgtt gttttggtat agtttggggt ttggtagtgg tgggtggggt       60
attggttaag agttgttatt gttgtggggga ggggagtttg gtttgttggg attgtaggta     120
atgggttgtg gggttttgtg ggttaggagg ggaatggggt tgggtgggtg agtagtgggt      180
agggggagttt agggtttggt tttgggtttt gttgttggat ttgggggttg tgaggaagag      240
ttgtgagttg agggtttggg gttggtgtat tttttttgtt ttgtttgtag ttggaaaatt       300
tttttttaag tttggggtgg tggagttttg ggggagaagg ggttgggggga gttgtggagg      360
gaggtgttgg gtttgtgtgt gtagggttta ggttgaggtt gggatgtggg tggggtgtag      420
gtttgggtta gggttgtagt tggttgtgtg ttgtgtttgt ttggggtgtt gttttttttt      480
tttttgggga gttgtgtggt tttttttttt tttttatttg tttttgtttt tagtttttttg     540
ttttgatata atgttttttt tgtgttgggt ttggttttttg tgttttgttt gttatggtag     600
ttgttgtttt tgtttttttgt gtggttgttg tttgggtttt gattgagggt tgatagtttt      660
```

```
tggttagggt ggtgttaggg tgggtattgt gttttttttt tttgtattat ttttttttaat    720
tggggtaatt ttttttgagg tgggaggtgt tggtttttttg gttttttttt tttttatttg    780
ggtaaagttt tttgtttttga atgatttttt ttgaagtgga tattttattt aaattgggta    840
attgttttta aaagggttat tgtgtttgaa tagttttttt tttggaagtt ttagtattta    900
gttaggtgtt ttggggtgtg taggttgttt tggtttttttt tttattggtg gttgtttatt    960
ttttgttttt tttttttggtt tgggtgggtt ggtttgggtt tttattttag agggtagttg   1020
gtttttttgtt ggtgtttagg ttgtagggtt gatgttttttg tttagttgag ggaaggggaa   1080
gtggagggga gaagtgttgg gttggggtta ggtggttagg gtgttgtatg gttttttttt   1140
ggttggtgtg tgttttttgta ggagagtgtg ttgggtagat gatgttggat atgatggagg   1200
tgtttggtta ttttaggtag ttgttgttgt agtttaataa ttagtgtatt aagggttttt   1260
tgtgtgatgt gattattgtg gtgtagaatg ttttttttttg tgtgtataag aatgtgttgg   1320
tggttagtag tgtttatttt aagtttttgg tggtgtatga taatttgttt aatttggatt   1380
atgatatggt gagtttggtt gtgttttgtt tggtgttgga ttttatttat attggttgtt   1440
tggttgatgg tgtagaggtg gttgtggttg tggttgtggt tttggggggtt gagttgagtt   1500
tgggtgttgt gttggttgtt gttagttatt tgtagatttt tgattttgtg gtgttgtgta   1560
agaaatgttt taagtgttat ggtaagtatt gttatttgtg gggtggtggt ggtggtggtg   1620
gtggttatgt gttttatggt tggttgggtt ggggtttgtg ggttgttatg ttggttattt   1680
aggtttgtta tttgtttttta gttgggtttt tgttgttgtt tgttgtggag ttgttttttgg   1740
gtttagaggt tgtggttaat atgtattgtg ttgagttgta tgtgttggga tttggtttgg   1800
ttgttgtatt ttgtgttttg gagtgttgtt gtttttttttt ttgtggtttg gatttgttta   1860
agaagagttt gttgggtttt gtggtgttag agtggttgtt ggttgagtgt gagttgtttt   1920
tgtgtttgga tagtttttttt agtgttggtt ttgttgttta taaggagttg tttttttgttt   1980
tgttgttgtt gttgttgttg tttttttaga agttggagga ggttgtattg ttttttgatt   2040
tattttgtgg tggtagtggt agtttgggat ttgagtttttt tggttgtttt gatgggttta   2100
gtttttttta ttgttggatg aagtatgagt tgggtttggg tagttatggt gatgagttgg   2160
gttgggagtg tggttttttt agtgagtgtt gtgaagagtg tggtggggat gtggttgttt   2220
tgtttggggg gtttttgttt ggtttggtgt tgttgttgtg ttattttggt agtttggatg   2280
ggtttggtgt gggtggtgat ggtgatgatt ataagagtag tagtgaggag attggtagta   2340
gtgaggattt tagtttgttt ggtggttatt ttgagggtta tttatgtttg tatttggttt   2400
atggtgagtt tgagagtttt ggtgataatt tgtatgtgtg tatttttgtgt ggtaagggtt   2460
tttttagttt tgagtagttg aatgtgtatg tggaggttta tgtggaggag gaggaagtgt   2520
tgtatggtag ggttgaggtg gttgaagtgg ttgttggggt tgttggttta gggttttttt   2580
ttggaggtgg tggggataag gttgttgggg ttttgggtgg tttgggagag ttgttgtggt   2640
tttattgttg tgtgttgtgt gataagagtt ataaggattt ggttatgttg tggtagtatg   2700
agaagatgta ttggttgatt tggttttatt tatgtattat ttgtgggaag aagtttatgt   2760
agtgtgggat tatgatgtgt tatatgtgta gttatttggg ttttaagttt tttgtgtgtg   2820
atgtgtgtgg tatgtggttt atgtgttagt attgttttat ggagtatatg tgtatttatt   2880
tgggtgagaa gttttatgag tgttaggtgt gtggtggtaa gtttgtatag taatgtaatt   2940
ttattagtta tatgaagatg tatgttgtgg ggggtgtggt tggtgtggtt ggggtgttgg   3000
tgggtttggg ggggtttttt ggtgttttttg gttttgatgg taagggtaag tttgatttttt   3060
ttgagggtgt ttttgttgtg gtttgtttta tggttgagta gttgagtttg aagtagtagg   3120
ataaggtggt tgtggttgag ttgttggtgt agattatgta tttttttgtat gattttaagg   3180
tggtgttgga gagttttttat ttgttggtta agtttatggt tgagttgggt tttagttttg   3240
ataaggtggt tgaggtgttg agttagggtg tttatttggt ggttgggttt gatggttgga   3300
ttattgattg tttttttttt atttagagtg tttttttgtta gtttgttttg ttgttgttgt   3360
gtggttttgg tttgtatttt agggagtggt ggggggtggtg tgtagggttt attgtgtttg   3420
ggataattgt agtgttgtta tagtggtggt tttatttttt ggtggtttta tttggttttta   3480
ttgttttgtg ttttagtttg g                                             3501
```

<210> 17
<211> 3501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 17

```
ttgagttaag gtatgaagta gtgaggttag gtgaggttgt tgagaggtgg agttgttatt     60
gtggtgatgt tgtggttgtt ttgggtatag tgggtttttgt gtgttgtttt tgttgttttt    120
tggggtgtgg gttaggttg tgtagtagtg atagagtggg ttggtgaggg gtgttttagg    180
tgggagagaa atggttgatg gtttggttgt tgggtttggt tgttaggtga gtgttttggt    240
```

```
ttagtatttt ggttgttttg ttggggttga ggtttagttt ggttgtgaat ttggttagtg      300
ggtagaggtt ttttagtgtt attttggggt tgtgtaggaa gtgtgtggtt tgtgttagta      360
gtttggttgt ggttgttttg ttttgttgtt ttaggtttag ttgtttggtt gtgaggtgag      420
ttatagtaaa gatgtttttg gggaagttga gtttgttttt gttgttgggg ttggggatgt      480
tggggagttt ttttaagttt gttagtgttt tggttgtgtt ggttgtgttt tttatggtgt      540
gtattttat gtggttgatg aggttgtgtt gttgtgtgaa tttgttgttg tatatttggt      600
atttgtaggg tttttttgttt gagtggatgt gtatgtgttt tgtgaggtgg tattggtgtg      660
tgaattgtat gttgtatgtg ttgtatgtga agggtttgag gtttaggtgg ttgtgtatgt      720
ggtgtgttat ggttttatgt tgtgtgaatt ttttttttgta gatggtgtat gggtagggtt      780
gggttagtta gtgtgttttt ttgtgttgtt gtagtgtggt tgggttttttg tagttttttgt      840
tgtatgatgt gtagtggtag ggttgtagta gtttttttag gttatttgga gttttggtga      900
ttttgttttt gttgttttta aaaggggtt ttaggttggt ggttttagtg gttattttgg      960
ttgttttggt tttgttgtat agtgtttttt tttttttat gtgagttttt atgtgtgtgt     1020
ttagttgttt agagttgggg aagtttttgt tgtatggaat gtatatgtat aggttgttat     1080
tgaagttttt gggtttgtta taggttaggt gtgggtatgg gtagtttttg aggtggttgt     1140
taggtggggtt ggggtttttg ttgttattgg ttttttttgtt gttgtttttg tagttgttgt     1200
tgttgttgtt tgtgttgggt ttgtttaggt tgttagggta gtgtggtggt ggtgttaggt     1260
tgagtggggg tttttttgggt gagatggttg tgttttttatt atgtttttttg tagtgtttgt     1320
tgggggagtt gtgttttttgg tttagtttgt tgttatagtt atttaggttt ggtttgtgtt     1380
ttatttagtg atagaggaga ttaggtttgt tggggtggtt ggggggtttg ggttttgggt     1440
tgttgttgtt gttgtgaaat gggttggaag gtggtgtggt tttttttttagt ttttggaagg     1500
gtagtggtgg tagtgatggt agggtgagag gtggtttttt gtaggtggtg gggttggtgt     1560
tgggagggtt gtttgggtgt gggggtagtt tgtgtttagt tagtggttgt tttggtgttg     1620
tggagtttgg tgggtttttt ttggataggt ttaggttata aagaggggag tagtggtgtt     1680
ttgaggtata gagtgtggtg gttgggttgg gttttgatgt gtatagtttg gtgtagtgtg     1740
tgttgattgt ggttttttggg tttgagggtg gttttgtggt aggtggtggt ggaggtttga     1800
ttggggatgg gtagtaggtt tggatgattg gtgtggtggt ttgtaggttt tggtttggtt     1860
gattataggg tgtgtagttg ttgttgttgt tgttgttgtt ttgtaggtgg tagtatttgt     1920
tgtggtgttt gaggtgtttt ttgtatagtg ttatgaggtt ggggatttgt aggtagttgg     1980
tggtggttag tatggtgttt aggtttggtt tagttttttgg ggttatggtt gtggttgtag     2040
ttgttttttgt gttgttagtt aggtggttgg tgtagatgaa gtttagtatt aggtggaata     2100
tggttgggtt tattatgtta tggtttaggt tgagtaggtt gttatgtatt attagggatt     2160
tgaggtaggt gttgttggtt gttagtatgt ttttgtgtgt gtggaagagg gtgtttttgta     2220
ttatgatgat tatgttgtat aagaagtttt tggtgtgttg gttgttgagt tgtagtagta     2280
gttgtttgga gtggttgggt gtttttattg tgtttagtat tgtttgttta gtatattttt     2340
ttgtggggat atatattggt tgggtgagag ttgtgtggtg ttttggttgt ttggttttag     2400
tttggtattt ttttttttttta tttttttttt tttttagttga gtgggggtat tagttttgtg     2460
gtttgggtat tggtgaagga ttggttgttt tttggagtgg gagtttaggt tggtttgttt     2520
ggattaggag aaggagtagg aggtgagtgg ttgttggtgg aggggaggtt agggtggttt     2580
gtatgtttta gggtatttgg ttgggtgttg gggttttttga gaagaaaatt gtttaggtgt     2640
agtgattttt ttggagatag ttatttgatt taagtaaaat gtttgtttta ggaaaagtta     2700
tttagggtgg agaattttat ttaagtaggg agaaagggag ttgaggaatt agtgttttttt     2760
gttttgggag aagttgtttt agttggggga agtgatatgg aggaggggag tgtggtgttt     2820
gttttggtgt tgttttggtt ggggggttgtt aatttttggt tggggtttgg gtggtggttg     2880
tgtggggagt ggaggtagtg gttgttgtgg tgggtagagt gtgaaggttg ggtttggtgt     2940
ggggagggtg ttatattggg gtaggaggtt gaggtaggaa gtaggtgggg gggagggggg     3000
agttatgtag ttttttaggggg aggggagggggg tagtgttttg ggtgggtatg gtgtatagtt     3060
ggttgtggtt ttgatttggg tttgtgtttt atttgtgttt tggttttggt ttgggttttta     3120
tatgtgtggg tttggtgttt tttttttgtgg ttttttttggt tttttttttt ttggaattttt     3180
gttgttttaa atttggggaa aagttttttta attgtagata gggtgggagg agtgtgttgg     3240
ttttaggttt ttggtttgta gtttttttttt gtggttttta aatttggtgg tagagtttgg     3300
agttgagttt tgagtttttt tgtttgttgt ttgtttgttt gattttgttt tttttttggt     3360
ttgtggggtt ttgtggtttg ttatttgtgg ttttggtggg ttggtttttt tttttttgtgg     3420
tggtggtagt ttttagttga tgttttatt gttgttgtta ggttttgagt tgtgttaggg     3480
tagtgttttt gttagttttg t                                              3501
```

<210> 18
<211> 2501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 18

```
tttttatagt gtaaatgtgt ttttattatt ttttggagta attttatttta aaattgtttt    60
tagtataaaa tttaaatatt taaatatgat tttgttggtt ttgttttttgt ggttttatttt   120
ttttttttttt taaatttagt tagtgtttgt gttgttgta atgttttttt tttttttgtag    180
gggttgttat tttaggtttt ggttttttttt tagaaagttt ttttttttttt ttttttagtgg   240
ggatagggtt tgtttatttt gatattatta gtttatttat atatattggt tataagttta    300
ggttgtattg ttattgaaag tttattattt gattttgagt agtttgagga ttttattaaa    360
atttaggaga tgtttagtaa atgttgattg aattatgatt gttttttaata tataaatgta   420
agattattta ggaatatttg ttaaaatgtt tttgttttttt gagattttat tttgggaggt    480
aagtagtggg ggtttaggat tttgtatttt gatagttttt tgatgtttgt atgtagaagt    540
gtagggatta ttatattgat aaatttttat tattttttaag ggggattttt tttttttaggg    600
gttattttttg gaagttttttt aaggataggg gttgtatgtt gttttttttag gttagtaatt   660
aaatttagaa aatgtttatt gagtgaatga tgaaatgata ggtgaataga tgaatgtaag    720
gtgttgagtt aattatttt ttatataagt tttagtagtt tttattgttt ttagttgtag    780
aaatggtttt tggaaggtaa gtttttgtgt atttttttta ttagaaattg gtaagttatt    840
ggattttaag ggagttgtgt gttttgtgtt tatttttttta ttagaaattg gtaagttatt    900
gattttttgtt ttgttttttgt tatttttttgt ttttttttgt tttgtagttg gtgtttagtg    960
gttttgtttg tttgtgtgtg tgttgttgta ggttttatt atgggtttat tgttgaggtt   1020
tgatgggtgg gtggtattgg ttattggtgt gggggtaggt gagtatgtga aggttggagg   1080
ttgtgttttt tgttgaggtg tagttggttg tttttttttgg gttggtatt gtgtgtagtt   1140
gtagttgagg ttatttgttt gaggtggtgg ggaggggaat ggttattttt gaggtattgt   1200
attttttgag gaggaaagag ttggaaatat ttggttttttt aagtaggtat agtttgtttt   1260
tttttagtat tttggtgtgg gttttttttaag gttttgtttg agaggagagg ttaggttggg   1320
ttgttgattg taaaattggg tgaaagttttt ttttgatttt tatttgtggg tattgattgt   1380
tattttttttt gtaattaatt ttttttagatt tttgtttagt tttttaaagg attgaaaagt   1440
tgtgagggg gggggttgga atttgttttt tgaagtgtag agatgttagt ttttgaaaag    1500
ttatttggtt gtttagtgtt tgtttttttt tgttgtaaga ttttaagttt gtgagaggat   1560
ttttttttaaa gagggtgttt gataagagtt tttttttttgtt ggagtttgta tgtttagtaa   1620
gttataattt gttttttgaaa tttattggag ttttggtaga ggttgtaagt ttaaatgtgt   1680
atagggggtta ggtgtatgat ggagaaagaa aatgggagta ggatgggtat atttgaggaa   1740
ttggagagta gagaattttg aagtggattg gttagtggga aagttgtttg tattttagga   1800
gtggtaaaat ggaaaattgt tatgtgaaat agttttattt tttaaagtat aaaaaattaa   1860
aataaattat ttatattaat atagatgttg tgtagtgaga ttttatatta gtttttttatt   1920
agtgggtgat ttttgtaatt tttaagtgta gggattttga tattatgtat ttttgatttt   1980
ttattggtag tatttttatat ttggaaaggt tttaatgtat gaattatttg agttatatat   2040
taaatgttat aaaattggaat tttgttaatt aatttttatg tatttttata tttgtattga   2100
taaagtggtt ttttatgttg ttttttagaa aatgttttta gtgttgatga atagttaagt   2160
attttatatt tatagttgtt tggttatttt tgtatgggta tgtatttggg tgtagttata   2220
ttttttaaat gttttttagga aaatatttttg tttatatttt gtttttattg taaataatgt   2280
attttataat gtttggtgtt ttaaattttt tttgatagtt tttggataat ttttatgtag   2340
gaggtttagg gattatattt taagatgttt ttgttattgt taaggagatt ttttttttta   2400
gggggttatat ttgaaaatta tttaaggata gggattgttt ttttttgatat tattagtata   2460
tttatatatg gtatgtagta tattttatat tagtatttag t                        2501
```

<210> 19
<211> 2501
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 19

```
attgagtatt ggtgtaaaat gtattgtata ttatgtgtaa gtatgttaat ggtgttaaaa    60
gaagtagttt ttattttttga atgattttta gatatagttt ttgaaaagga aagttttttt   120
agtgatggta aaaatgtttt agaatgtaat ttttggattt tttgtatgaa aattatttaa   180
gagttgttaa aaaagattta aaatattaag tgttgtaaaa tatattattt ataataaaag   240
taaagtgtaa ataaaatgtt ttttttaaaaa tatttagaag gtatgattat atttaaatat   300
atgtttatgt agagataatt agatagttat gggtataaaa tatttggtta tttattaata   360
ttgaaagtat tttttgaaag gtagtataag aagttatttt attaatatag atatgaaagt   420
```

```
atataggaat taattgatag aattttagtt tgtaatgttt aatatataat ttaaataatt    480
tatgtattag ggttttttta ggtataaggt attattagtg gagaattaaa ggtgtataat    540
gttaagattt ttgtatttgg aggttataga ggttatttat tggtgagaaa ttaatgtaaa    600
attttattgt atagtattta tgttggtatg aatggtttgt tttaattttt tgtattttaa    660
aaaatggggt tattttatat aataattttt tattttgttg tttttgaaat ataggtaatt    720
ttttattgg ttggtttatt ttggaatttt ttgttttta gttttttaga tgtgtttatt    780
ttatttttat ttttttttt tattatatgt ttgattttg tgtgtatttg agtttataat    840
ttttgttaag attttagtgg attttgagaa tagattgtga tttgttaagt atataaattt    900
taatgggaa gggttttat tagatgtttt ttttaaagaa ggtttttta tgaatttaaa    960
attttatgat agagggaaat aaatattgaa tgattgaatg attttttagg agttgatatt   1020
tttgtgtttt aggggtgaa tttagttt tgtttttgt ggttttag tttttaaaa   1080
gattaggtaa agatttaaga gagttaattg taggaagagt aataattgat gtttatagat   1140
aagggttagg gagaatttt atttagtttt gtaattagta gtttagttg gttttttttt   1200
ttaggtagga ttttgggaag tttatattgg ggtgttgggg agaagtgggg tgtatttgtt   1260
tgagagatta ggtgtttttg gttttttttt tttaagaga tgtggtgttt taagaataat   1320
tattttttt tttattattt tagtgggggtg atttttagttg tggttgtgtg tgtatgttgg   1380
tttgaaaaga gtagttagtt gtgtttttagt aaggggtgtg gtttttaatt tttgtatgtt   1440
tatttgtttt tgtgttggtg attagtatta tttgtttgtt gaattttagt ggtgagttta   1500
tgaataaggt ttgtaatgat atatatatga ataagtagag ttgttggatg ttgattgtgg   1560
gataggagga ggtgggggaat ggtgggggtg ggatgagggt tagtgatttg ttgatttttg   1620
gtaggaagat gagtgtagag tgtgtggttt ttttggaatt ttgggaaatg tagttaagag   1680
tgattttatt tgttggaaga tttgtttttt aggggttatt tttgtggttg gaagtaatgg   1740
gagttgttag gatttgtgta gaagaatagt taatttgata tttttgtgttt atttatttat   1800
ttgttgtttt attatttatt taataaatgt tttttgggtt tagttgttga tttagagaaa   1860
tagtatgtgg tttttatttt tgaggggttt ttagagatag ttttgggaa ggaaagtttt   1920
ttttagggat ggtaaagatt tgttagtgta ataattttg tatttttata tgtaaatatt   1980
agggggattgt taaaatgtag agttttggat ttttattgtt tattttttaa aatagaattt   2040
taaggggtaa aaatattttg ataagtgttt ttaaatgatt ttgtgtttgt atgttgagaa   2100
tagttatagt ttaattaata tttattgagt attttttgag ttttgatagg atttttaagt   2160
tatttagagt tagatggtaa atttttaata atggtgtagt ttagatttgt gattaatgtg   2220
tgtaagtgag ttaatggtgt taaaataaat agatttatt tttgttgggg agaaagagga   2280
aaaatttttt gaaggaggat taggatttaa agtggtgatt tttgtaaaga aagaagggta   2340
ttataggtag tataaatatt agttaggttt ggggagaaag agggtaaagt tataaaagta   2400
aagttagtaa gattatgttt agatgtttga attttgtgtt gaaaatggtt ttaagtagga   2460
ttattttaga gagtggtggg aatatattta tattatggaa a                       2501
```

```
<210> 20
<211> 2470
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 20
```

```
aaagatgatt aaaagttttaa ttgtttattt gaagagttga tttttttatt tttgtaataa    60
agggtatttt tagtagtttt tgtttatttt gtttatttgg ttttttttgt ggttgtgtaa   120
ggttataatt tttgtgtttt agtaaatttg tgtatgttta tttttttttt tgttattatt   180
tttttttta ttttgttta ttatttgat gtaaaattat ttgttaattt tatttgaaat   240
gagaaatttt aaggtttata ttatttaaat tttgttagat tttatttttt gttatatggt   300
ttataatgtg ttgggtattt ttagatttgt ttattaaaaa gatgtaaaat aaaataatga   360
ttattttgt ggattttttt tttatttttg agatgttttt tttggttgta ttatttttt   420
attttttgtt tattgattag aggaggggtt ttaattatgg gtgaatttta tattttattg   480
aagaggttat gttatatgta tatttttata atataattta tatttatata gtattttat   540
tttttagtata ttttttttta ttaatttaa taatattatt gtaagttatg ttgaagtaga   600
ttgtaagtgt ttatttataa attgtgaaat gaattaaaat gaaagggtaa agattaaatt   660
atgattaggt ttgaaattaa tatataagat ttaattttt ttaattaaag atttttgtag   720
gtgatttttg tttgtaggat tttttttttt tttttagatgt tattggattg tattaggttt   780
attgtagatt ttagttgtt tagaattaat tagatttaag atgagttttt tgattttttt   840
tggtagagtt ttttaattgt tgaatttttaa tattgttgtg attagttagt gttataattt   900
gtttgtttta ttttgtgtaa tggatttttat attatagagg tattttttta atgttaagat   960
gtttaagtat tgtttaagtg taaattattt aatatttttt agttattaag taattaagat  1020
```

```
aggtaggatt ttatttgttt taaaatgatt tgatttaaat taaaaagaga atgtggattt    1080
tttgaatttt atttggttaa ttttaatata atttttagta ttttataatt ttttttaaag    1140
tttttttatt tggttatttt ttgtattttt tttgtttttt ttttttttttt ttagttataa    1200
taattgttag attttgtttt atttttttttt gatagttttt atttttaagg ttatttattt    1260
tttttaggta tttttttggtt ttagtttgag tatagtagat tttaagatta tatatgttat    1320
agtataggtt attatagtta atttttttgaa taaatgtgat tgaattttat gttagtaatt    1380
tttatttatt attttttttat taaaaaggtt taaagttttt atttaatgtt ttttttttatg    1440
tttatttgt taaatgattg tttttttaatg atattttaga atttttagaat tattttttatta    1500
tggaggatgt gtaagattag tttttttatta aataaaaagt gtgaaatgga atatgtaatt    1560
ttattaattt attttggttt taaaattttg tgattattag ataaaattta gaaataaaat    1620
agtattatta atataaataa attttttatta taattatatt ttttaagttt tgtttgtaag    1680
aatgggtaaa atatttttaa aattttgaag aaattattat ttgatagaaa gtttaattta    1740
tttgtgagaa ggtaaatgta tttagatata attaaagttt tttttttttat tttaatttta    1800
tttatttttga attaagattt tattgtttta tttttttaga tgttgttatt tgaataatat    1860
tgttttgaga ttaaaaatta gtatattaat ataattttttt ttaaatgttt taagagtttt    1920
gtttttttttta tttttttttt taaaaataag tagttattaa attttttagt agtgaatttt    1980
aaaattttttt ttaatttttat aggtttaagg gtagttaagg atggttgtag ttttatatga    2040
ttagttgtta aagtaagttg aggtattgaa gatggagaat ttaaatttttt gataagagtt    2100
agaagataat tttaattttat ttataaaaatt ggaaattgag gtatttaata tgaaggtatt    2160
aagattgtga ttttttaattg tagtttatttt attttttattt agtattttttt tttgtaaatt    2220
tgaggtaaga tattttttatttt aaaagtgtat tttaaattaa gtaataatat gtaaattttt    2280
ttttgtaaaa gttagtatttt atatttttttaa ataagatata ttgaatttat ttagtgaatt    2340
atataaagaa aataagtgta aaatttttaat ggttagttag ttttttagttt tttttttaagat    2400
taaagagaag agattaaata tagtattatt gtattgaggt aaggttttttt gtgtagttta    2460
                                                                      2470
tagaaattag
```

<210> 21
<211> 2470
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 21

```
ttagttttta tgaattatat agaaaatttt gttttagtat agtgatgtta tatttggttt      60
ttttttttttta attttaaaaa gaattaagaa ttaattagtt attggagttt tatatttatt    120
tttttttatat gatttattga atgaatttaa tatattttat ttaaaaatat aaatgttaat    180
ttttgtaaga aagagtttat atattattgt ttaatttaaa atatattttt aagtaaagtg    240
ttttatttta agtttataag agggaatatt gaataaaaat ggataaatta taattaaaag    300
ttatagtttt gatattttta tattagatgt tttagttttt agttttgtaa gatgattgga    360
attatttttt agttttttgtt gaagatttga gttttttatt tttagtgttt taatttgttt    420
taataattga ttatatgaag ttgtagttat tttttggttat ataaggttaa                480
aaaggatttt gaaatttatt attaaaaaat ttagtggttg tttgtttttta aagaaagggg    540
taaaggaaat aaaatttttta agatgtttaa gaagaattgt gttaatatgt tagttttttgg    600
ttttaaaata atattgttta agtagtagta tttaagagga tgaaatagtg gagtttttagt    660
ttaagataaa tgaaattaaa atagaagaga gaatttttagt tgtgtttgaa tatatttgtt    720
tttttatatga tggattaaat tttttatttaa gtaataattt ttttaaggtt ttaaagatat    780
tttatttatt tttataggta aaatttagga aatataatta tgataaaaat ttatttatat    840
tagtaatatt attttatttt tgaatttttat ttgatagtta tagaatttta gagttagaat    900
ggattaatga gattatatat tttattttat atttttttatt tgataaaagg ttaatttttat    960
atattttta tggtgaaata gttttgaagt tttaagatgt tattaaaagg taattatttta   1020
ataaaatgga tatgaaggag agtattaaat gaagatttta agtttttttg ataggaagat   1080
ggtaaataag aattattaat ataaagttta attatattta tttaaaaggt tgattataat   1140
agtttatgtt atggtatatg tggttttggg atttgttgtg tttaaattga ggttaaaaga   1200
tatttaaaga gaatggatga ttttaggagt agagattgtt aaagagaaat gaagtagagt   1260
ttggtagtta ttatgattgg gaaagaagag gagagataaa gaagatataa aagatagtta   1320
ggtaagagga ttttaggaag aattatagaa tgttaggagt tatattaaga ttaattaagt   1380
aagatttagg agatttatat tttttttttta gttaggtta aattattttg gaataaataa   1440
aattttgttt attttaatta tttaatagtt aaaagtatt aagtagtttg tatttaagta   1500
atatttaaat attttgatat taaaaaaatg ttttttgtaat atgaaattta ttatataaaa   1560
taaggtagat aggttgtaat attggttagt tatgataata ttggagttta gtaattggaa   1620
```

```
gattttatta aaggaaatta ggggatttat tttagattta gttagtttta taatggttag       1680
aatttatagt aaatttggta taatttaatg atatttgagg aggaagggga gttttgtagg       1740
tagggattat ttataaaagt ttttggttga aaaaaattga gttttgtgtg ttaattttag       1800
gtttggttat gatttaattt ttgttttttt attttaattt attttataat ttgtaaatga       1860
atatttataa tttgttttaa tataatttat agtgatatta ttaggattaa taaaaaaagg       1920
tatgttaaaa ataaaagtat tatgtaaatg taagttatat tatgaaaata tatatgtaat       1980
ataatttttt tagtaagata tagggtttat ttatagttaa gatttttttt ttgattaatg       2040
ggtaaggggt gaagaagtaa tgtagttaaa ggagatattt taaaaataaa ggaaaaattt       2100
ataggagtga ttattatttt gttttatatt ttttaataa gtaggtttga aaatatttag        2160
tatattataa attatatgat agaggtaggg atttgataga atttgaataa tgtgaatttt       2220
aaaatttttt attttaaata aaattaatag gtaattttat attaaaataa taaaataaaa       2280
taagagaaaa ggtagtaata gagaaaaaaa tgggtatgta taagtttatt gagatatada       2340
agttataatt ttatataatt ataaaaagag ttggatgggt aagatgagta gagattgtta       2400
aaagtatttt ttattatagg aataaaaaaa ttaatttttt agatgaataa ttaaattttt       2460
aattattttt                                                             2470


<210> 22
<211> 7001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 22

aatgtaatgg aaaaagagag attgtaaagt tagaaggttt aggaattgtt ttttgattag         60
gtgtggaagg taagggaaaa ttagtttttg aagaagatag tgagattta atttgggtgg         120
ttggagagat agtgatgttg ggtatagata tggggaagtt gagaggaata ttatgtttga        180
gaatggtgat ttatatttga ataagtttgt aatgtttagt agattgttgg aaaagtgggg        240
ttggagatat atttaatgga ggagttagat taattttat ttttttttat ttgagagagt         300
tagtaagtta tggttggaat gtgtgtgttt agtaggagag ggtagggagg gaagttaaga       360
gagttgggag tttgagtgaa gttttttgtta aaggtagaag aggaaagttg gtgtagtata       420
gtatatttt ttatttatgt ttattaagtt tagggataag gtttattaag atgagtttgg        480
aagagaatgt tggagagaaa gtggttaaga aaattgtttt tattgaattt tttgggttaa        540
ttttgattgt aagttttga ataattaaag tttgtgagga gatagttaat ttttttattt        600
ttttatgtt aatagtgaat aattgtagat tttttttttt tttttttttt ttttttttgt        660
ttttttttt tttttttttg aatatttttg tttttttttg ggattggttt agagtatggg        720
tggttattgt tgatttatag gaggtattat tgttaatattg attatggaga gagtttttt       780
tttaatattt atttatattt agtatttatt tttaatattg attatggaga gagtttttt        840
gtgtttaaat attgtaatat tgggggtttt ttaaagtata aaaatatata tttgtatgat        900
ggtattatta atatttttat ggttttttat ttttttttg tattggtttt aagagttatt        960
tataaatttt ttagtaattg tatagtgttt tagggttaga gattggttat ttttggtatt       1020
gtgattagag ttatttaata tttaaggtgg tgattaatgt ttggtaataa agttttattt       1080
gggtgttatg tgtttttggga ttttgagtgt gggtatttta ggagtatttt agtattgtgt      1140
gttagtatta tggttgagag aatagttgag aaagtggtta agaggtggat ttatgtgaat       1200
gttattggga aatgagagat tttgttttta attatggtta gtgtaatttg aaagtttaaa       1260
attagtttaa aataaaggta tttatttta ttttatgttt atatttagg ttttaataa         1320
tatgtatttt ttatatgttt atagaaagta gttaattgag ttatttatgg aaaggtttgt      1380
gggtttggtt aatgaagtgg aggagtatta tattttagtt ggaaatatat ttttagaatg      1440
ttaaaatatt tattttaaag tttggttttt tggtgtaatt ggaggtatgg taatgttttt      1500
gtttagagat tgggggttag ggttagtaag gtatttgatt tatatgtatt ttagaaggtt      1560
tttattgtta aattatattt ttttggaaaa attatttatg ttttattttg taaatttgat      1620
atttatatat ttttgattgg tatttattt tagttgtaag attatgattt atagtaagtt       1680
tgtttttttt tttgtttggg gtggtagtag aaagtatagg gtatttttta gttttaagg      1740
gtagggtaa aggggttggg gttttttttt tttagtatag tttttttgg ttgtgttata        1800
ttgttttttg tgagtagata gtaagttttt tttttttttt tattgttatt tatttagtgt      1860
tgtgtagtag tttagttgtg tgtttgttgg gaggggttgt taagtgtttt gtttattggt      1920
tgtttttga attttgtta ttttatgtat aaatatattt atatttttt tttgtttagt        1980
ttatatattg agttatttgt atatgtgagt atatttttt tttttttttt attttttttgg      2040
ttttgattt ttataagttt atggaatatt tttggaaaga tgtttttgat ttagtagggt       2100
aggtttgttt tgatttttt ttttgtagtt ttagtatttt gagaaagtaa tttatttttt       2160
tggttagtgt ttgtatttta gtagggagat gaggattgtt gtttttatg ggggtatgtg      2220
```

```
tgtgttttt  tttttttta  ggatttgtag  gattttttgt  gttatttgta  tataatttgg   2280
taggtttata  tttttaaga  gttttatgaa  gtgttttttg  tatgtgtttt  aaaaaggtat   2340
ttgaaaattg  aaagtgtgat  ttatggaaat  taaattattt  gtaaaaaatt  gttttggaaa   2400
gtaatgattg  ttggttataa  agggaaatat  ttgtgatgta  tttaatgtgt  ttttaatttt   2460
ttatttgttg  ataatttata  gttattaatg  ttaaatttga  ttttggtttt  agttatattt   2520
gtatattgtt  taataatggt  ttattttgt   aagaattaga  taaaatgtat  atttgatata   2580
aaatagttaa  aaatgtaatt  tttagtaata  gtaagtttgg  tatttagata  gattatgaat   2640
attttgttag  atattttgtt  gggtgtttgg  gatagtaatt  aaaataaagt  attgatagtt   2700
gtattagagt  ttattaggtt  gtagtaaagg  aagtttattt  aaaagtataa  attatttaag   2760
attatagatg  tatgatatat  tttatttatt  ttttgttttt  ttaatatgta  tatatatata   2820
tatatatata  tatatatata  tatatgtgtg  tgtgtatgtg  tgtgtgtatg  tttaattttt   2880
aatttagtta  aaaatttttt  tttatttgtt  ttttatttgg  atatttgatt  ttgtatattt   2940
tagtttaagt  gaattgagaa  gattgagttg  taggattaaa  ggatagatat  gtagaaatgt   3000
atttttaaaaa  tttgttagtt  ggattagatt  gataatgtaa  tataattgtt  aaagtttggg   3060
tttgtgattt  gaggttatgt  ttggtatgaa  aaggttatat  tttatattta  gtttttttgaa   3120
gttttggttg  tataattaat  ttgtggaagg  tatgaatatt  tatgtgtgtt  ttaattaaag   3180
gttttttttga  attattttt   atatgagaat  ttttaatggg  attaagtata  gtattgtggt   3240
ttaatataaa  tatataagtt  aggttgagag  aatttagaa   ggttgtggaa  gggtttattt   3300
attttgggag  tattttgtag  aggaagaaat  tgaggttttg  gtaggttgta  ttttttgat    3360
ggtaaaatgt  agttttttttt  atatgtatat  tttgaattt   tgttttttttt  tttttagatg   3420
ttttttgtta  gttttttttag  ttgttaaata  tagttgtttg  tggttggttg  tgtatgtaat   3480
tgtatatttt  attttatttg  ttttattttg  gttatagtgt  agttttttt   agggttattt    3540
tatgtatata  ttatgtattt  ttagttaatg  aggaggggga  attaaataga  aagagagata   3600
aatagagata  tattggagtt  tggtatgggg  tatataaggt  agtatattag  agaaagttgg   3660
ttttttggatt  tgttttttgt  gtttattta   agtttagttt  tttttgggtt  attttttagta   3720
gattttttgtg  tgttttttgtt  ttttggttgt  gaaatttagt  ttttatttag  tagtgatgat   3780
aagtaaagta  aagtttaggg  aagttgtttt  ttgggattgt  tttaaattga  gttgtgtttg   3840
gagtgatgtt  taagttaatg  ttagggtaag  gtaatagttt  ttggttgttt  tttagtattt   3900
ttgtaatgta  tatgagtttg  ggagattagt  atttaaagtt  ggaggtttgg  gagtttagga   3960
gttggtggag  ggtgtttgtt  ttgggattgt  atttgttttt  gttgggttgt  ttggttttat    4020
tggatttgta  ggttttttggg  gtagggttgg  ggttagagtt  tgtgtgttgg  tgggatatgt   4080
gttgtgttgt  ttttaatttt  gggttgtgtt  ttttttttag  gtggtttgtt  ggttttttgag   4140
ttttttgttt  tgtggggata  tggtttgtat  tttgtttgtg  gttatggatt  atgattatga   4200
ttttttatat  taaagtattt  gggatggttt  tattgtatta  gatttaaggg  aatgagttgg   4260
agttttttgaa  ttgtttgtag  tttaagattt  ttttggagtg  gttttttgggt  gaggtgtatt   4320
tggatagtag  taagtttgtt  gtgtataatt  attttgaggg  tgttgtttat  gagtttaatg   4380
ttgtggttgt  tgttaatgtg  taggtttatg  gttagattgg  ttttttttttat  ggttttgggt   4440
ttgaggttgt  ggtgtttggt  tttaatggtt  tgggggggttt  ttttttattt  aatagtgtgt   4500
ttttgagttt  gttgatgtta  ttgtatttgt  tgttgtagtt  gttgttttttt  ttgtagtttt   4560
atggttagta  ggtgtttttat  tatttggaga  atgagtttag  tggttatatg  gtgtgtgagg   4620
ttggtttgtt  ggtatttttat  aggtatttgt  gtttgtgttg  tttgttgggg  tggttgttgt   4680
gtttggtagg  agggaggggag  ggaggggaggg  agaagggaga  gtttaggag   ttgtgggagt    4740
tgtgggatgt  gtgatttgag  ggtgtgtgta  gggagtttgg  ggtgtgtggt  ttagtttggg   4800
ggttttgtgt  gtagtttgtg  ttgtgtttag  agtttaagttt  ttttgttggg  tagttgaaaa   4860
aaatgtattt  tttatttatt  tattgtttgt  gtgagaggta  gatttgaaag  tttgggtttt   4920
ttaataaaat  atatgttgga  aaattagata  aagtagtagt  tatttgtggg  ggaaaatatt   4980
tttaggtaaa  taaatatggg  gtgttttgag  ttatttggga  aggttttgtt  tttggtattt   5040
aaagttgggg  gtgtttggag  ttagtagagt  ttagtagagt  tttatttatt  tttttaatgt   5100
ttttgtttaa  tgtgttttttt  aaatttttttt  ttatttagat  tatttgattg  gaaatatgtt   5160
agttatgatg  atgatttttt  gggaagtgat  ttttgttatt  tgtttttttt  tttttttttat   5220
tttatgtttt  ggggttttag  agagtgattg  ggagttgaat  gggtttgatt  ttggagttag   5280
ttggttgagt  ttgtgttgga  gtggattgtt  ggtatgtgat  ttttgatagt  tggaaatttg   5340
taggtgtttt  gtgagtttaa  aataagttat  atggaagtat  aagtgtttaa  aaataatttt   5400
ttgttagttt  agtgataagt  ttgtttttatt  tggggagaat  gttttggagt  ggtgtgtggg   5460
ttagttaggg  tttgtgtttt  gtagttattg  tggaaggagt  gtggttggtt  taggatatag   5520
gagattattt  tgtgatttta  atggtgaagg  ttgtgtgttt  ttattttaat  tttttttttt    5580
ataagaattg  ttttttttttt  ttttttttttt  tttttattt   ttttttgttt  agtttttttt    5640
tttgttttttt  gttttttgtt  ttttgatgg   gtttgtagag  ggattaggtg  ggtgttttttg   5700
gtgaatattt  tttaggtgg   ttataggata  ggtgtatttt  ggattgggtt  tggaagtttt   5760
agggtgttat  atggttgggt  tttgaattag  gtatttttta  attgtatatt  ggtatttgga   5820
ttggtgtttt  tatattttttt  tgttttgtaa  gttgtggatt  agttttttgtt  tagtattttg   5880
ttttttaggga  tatttatagt  agaaggaagg  ggattaaagt  gtagtttggt  tttagaggat   5940
attgaagggt  agattttggg  ggtatttagt  gtgtattttt  agttgttttg  gagaaattta   6000
```

```
gagtatttta tagttatgta gatttaagtt gttttattt aaaagataaa taatgaataa    6060
aattttaaa ggttggtata ttttaaatta attttattg ttttaattta gggttaaaat     6120
agagaaaaag gattttttt gtttattttt tttttttaa atggaagaat aaagtatagt     6180
gattaagttt aattttatat aatatttaaa attgtttgat gtgaaggaag gtattggtat   6240
gatgtgaatt ttataatttt atgatggatt ttagaaatta tttttttttt tatttaattt   6300
ttagtttttt tattgtaaat taatgttgtt gaattttaat gggtattaat gagattgttt   6360
tttggtagat tatttattgt tttgttaata attataaagt gaatttggtt aaatatagag   6420
gggattgtat tttatttaaa attgtttatt attttagtga taagtggtat tagtgtaata   6480
tgttttattt tatatttttt gtattatatg atatttaaat atttttagaa taataaaaaa   6540
agagataagg aatttaaaaa ttaaaaaaaa aatttgtata aatgggattt tgtgtggaaa   6600
tttagtttta gaatgatttt ttttgtgttt tatttttggg attattttt tttttttgtt   6660
agaattttgt ttgttattat ttagtaagga aaagaagtat ttatgtaagt ttttatatg    6720
gatagatatt atttagtatt tttttttttt tagtttttt gtttaaatga ttttgggtat    6780
aaaggaaagg attgattggg tttttttagg aaattttaag tttttaagt agtttttaaa    6840
agttttgggg ttgaaagtag tgtttttaaa ttgtttgtta tgatttagag ggttatgaat   6900
ttagtttagt gagtttagaa tatttttaa aaggattaaa atggaaagga atataatagа   6960
aaatattaga gtgtatggta ttttgtaagg ataagttttg t                       7001
```

<210> 23
<211> 7001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 23

```
ataaaattta tttttatgaa atattatgta ttttgatatt tttattata tttttttta     60
ttttagtttt tttaaaaaat attttagatt tattaaattg agtttatgat tttttgggtt    120
atgataagta gtttgaaaat attgttttta gttttaaaat ttttgagaat tatttaagaa    180
atttaaagtt ttttaaaaga gtttaattaa tttttttttt tatatttaga gttatttaag    240
tagaaaaatt gagaggggaa aaatattaaa taatatttgt ttatatgaag aatttgtata    300
gatgttttt tttttttgttg gataataata ggtagaattt taataaaaga ggaaagataa     360
tttgggaaat aaaatatagg aaaaattatt ttaaaattga atttttatat agagttttat    420
ttgtgtaagt tttttttta atttttaagt tttttgtttt tttttttatt attttaagag     480
tgtttgaata ttatgtaatg tagaaagtgt aagatagggt atattatatt gatattattt    540
attattggga tgatgaataa ttttgaataa gatgtgattt tttttgtatt tgattaggtt    600
tattttgtaa ttattagtaa ggtagtaaat aatttattaa ggagtagttt tattagtgtt    660
tattgaaatt tagtagtatt aatttgtaat aaaagaattg aaaattaaat agggaagaaa    720
atggttttg gagtttatta taaggttatg gaatttatat tatattagtg tttttttta     780
tattaagtag ttttaaatgt tgtgtggaat tagatttaat tgttgtattt tgttttttta   840
tttaaaaaaa aaaaggtggg tagaagaaat ttttaaaagt tttgtttatt gtttgttttt   900
ataagtaaaa ttaatttgaa atatgttaat tttttttttt tgttttaatt ttaaattaaa   960
tgagtaaaga tagtttggat ttgtgtggtt gtgggatgtt ttaaattttt ttaaggtggt   1020
tgaagatgta tattgaatat ttttagaatt tgttttttag tattttttgg ggttaaattg   1080
tattttagtt tttttttttt tgttataaat attttttggaa atagaatatt gaataaaaat   1140
tggtttatgg tttataaggt agaaagatat agggatattа gtttggatat tagtgtatag   1200
ttgggaaatg tttaatttag gatttagtta tgtggtgttt tgaagttttt aaatttagtt   1260
tggggtatat ttgttttgtg gttatttagg aaggtgttta ttagaagtgt ttatttaatt   1320
tttttgtagg tttattagga aaataaaaaa taaaaataa aaggagaaat tgggtaagag    1380
aaaatgggag ggagaggaga gggagaaaga ataattttg tagggaaaaa aattaaaatg     1440
aggatatata atttttgtta ttgaagttat aaagtggttt tttgtgtttt ggattggttg    1500
tgtttttttt atagtggttg tgaggtgtag attttggttg atttgtatgt tattttgggg   1560
tattttttt gggtgggata ggtttgttat tgggttggta ggagattatt tttaagtatt   1620
tgtgttttta tatggtttgt tttaaatttg tgggatattt ataattttt ggttgttaga   1680
agttatatgt tagtaatttg ttttagtgtg gattagtta gttaattttg aaattagatt    1740
tatttaattt ttaattgttt tttaaagttt taggatgtgg ggtggggagg aggggaaagt   1800
gggtgatagg aattgttttt tagaaagtta ttattatagt tgatatattt ttaattaaat   1860
agtttagatg aaaggaaatt tggggagtat attaaataaa aatattaaaa ggataaataa   1920
aattttgttg agttttgtta attttaaata tttttaattt taaatgttaa gagtgagatt   1980
tttttaagtg atttaaagtg ttttgtgttt atttgtttgg aggtgttttt ttttataaat   2040
aattgttgtt ttgtttggtt ttttaatgtg tgttttgtta ggaagtttgg gtttttgggt   2100
```

```
ttgttttttg tatggatggt aagtgggtgg agagtatgtt ttttttagtt gtttggtgag   2160
agaatttgat tttgaatgta gtgtgggttg tatgtagaat ttttgggttg ggttgtgtgt   2220
tttgggtttt ttgtgtgtat ttttgggttg tgtgttttgt ggtttttgta gttttttagg   2280
tttttttttt tttttttttt tttttttttt ttttgttggg tgtggtggtt attttgatgg   2340
gtggtgtggg tgtgggtatt tgtagaatgt tggtgggttg gttttgtgta ttgtgtagtt   2400
gttgggtttg ttttttaggt agtagggtat ttgttggttg tggggttgta ggaaaggtga   2460
tagttgtggt ggtgggtgta gtagtattag tgggtttgga gatatgttgt tgagtggggg   2520
gaaatttttt aggttgttgg agttgaatgt tgtagtttta gatttggggt tgtaggggag   2580
gttggtttga ttgtagattt gtgtgttggt ggtggttgtg gtgttgaatt tgtaggtggt   2640
gttttgggg tagttgtata tggtgggttt gttgttgttt aggtatattt tgtttagggg   2700
ttgttttagg gggattttga gttgtggatg gtttagggg tttagtttgt tttttttggat   2760
ttgatgtagt agggttattt tagatgtttt ggtgtggagg gttatggtta tggtttgtgg   2820
ttgtgggtag ggtgtagatt gtgtttttgt agggtagaag gtttagaaat tggtgggtta   2880
tttggaaaaa gagtatagtt tgaggttaga ggtgatgtag tgtatgtttt gttgatatgt   2940
gagttttggt tttggttttg ttttgggagt ttgtgggttt ggtgaagttg ggtgatttga   3000
tgggagtaag tgtagtttta ggatgaatgt tttttgttag tttttgggtt tttgggtttt   3060
taattttaag tattggtttt ttgagtttat atgtattata aaggtgttgg aggatggtta   3120
gggattgttg ttttgttttg atattggttt aaatattatt ttaggtataa tttgatttgg   3180
agtgatttta aagagtagtt ttttttgaatt ttatttttatt tgttgttgtt gttggataga   3240
ggttgagttt tatggttagg gggtggggt gtatgaggat ttgttaaagg tggtttaggg   3300
aagattgggt ttaaaataaa tgtgaaagat ggatttaggg gttggttttt tttaatgtgt   3360
tgttttatgt gttttgtgtt agattttgat atatttttgt ttgttttttt ttttgtttga   3420
ttttttttttt ttgttggtta gaaatatgta gtgtgtatat aggatgattt tggggaggat   3480
tatattgtaa ttgagatagg gtagatagaa tggggtgtgt ggttgtatat gtagttagtt   3540
atagatagtt atatttagta gttggggga ttgataggg gtatttgagg ggaaggggt   3600
ggagatttag ggtatatata taggaagagt tgtattttgt tattaggaga atgtaatttg   3660
ttaggatttt agttttttttt tttgtaaaat gttttttaaag tagatagatt tttttataat   3720
tttttgagat ttttttagtt tgatttgtgt gtttatgttg gattatagta ttgtatttgg   3780
ttttattagg aattttttatg tgaaggatga tttagaaaaa tttttggtta gggtgtatat   3840
gggtgtttat gttttttata ggttggttat gtaattaaaa tttttagaaaa ttgaatataa   3900
aatgtgattt ttttatatta aatataattt taggttatga attaaagttt tggtaattat   3960
gttatattgt tggtttggtt tagttaatag attttttaaaa tgtatttttg tatgtttatt   4020
ttttagtttt ataatttgat ttttttttggtt tatttgggtt aggatatgta gaattaaata   4080
tttagatgaa aaataaatag aaaaaagttt ttaattgaat taaaagttaa atatgtatat   4140
gtatatatat atatatatat atgtgtatat atatatatat atatatatat atatatatta   4200
aggagataaa aaataggtga agtatattat gtgtttataa ttttggatag tttatatttt   4260
tgaataaatt tttttttgttg tagtttaata gattttgata taattattaa tattttgttt   4320
taattgttat tttaaatatt taatagagta tttgatgaag tgtttatggt ttatttaaat   4380
gttaagttta ttgttattaa gagttatatt tttgattatt ttatattaag tatatatttt   4440
atttaatttt tataaaaata gattattgtt ggataaatg taaatgtagt tgaagttaaa   4500
attgagttta gtattaatga ttatagattg ttagtaaata aagggttaaa aatatattag   4560
gtgtattgta gatatttttt tttatggtta gtaattatta tttttttaaag taattttttta   4620
tagatgattt aattttttata aattatattt ttaatttttta aatgtttttt taaaatatat   4680
gtaaaaagta ttttataggg tttttaaaaa atgtgaattt gttaaattat atgtaaatgg   4740
tataaagaat tttataagtt ttgaaagaaa aaggagatat atatatattt ttatggagaa   4800
tagtaatttt tatttttttg ttaggatata gatattagtt agaaaggtaa gttgtttttt   4860
taaaatgtta aagttataga gagagaaatt aaaataagtt tattttgttg gattaagaat   4920
gtttttttag aaatgttta tgggtttgta gaagttaagg gttgagagag tgagaaggaa   4980
ggaaggaatg tgtttgtatg tgtgagtggt ttagtgtgtg aattaggtag agagagtgtg   5040
tggatgtgtt tgtgtgtgga atggtaggga tttgggaagt agttagtagg tagggtattt   5100
ggtagttttt tttggtagat atgtttgttt atagggagta gtgtggtata gttagagaaa   5160
ggggagtgag gggagatttg ttgtttgttt tttgttttta tttttggagg ttggaaagta   5220
gttgtattgg ggaggagaaa tttagttttt tttgttgtta ttttaagtaa gaggaaaaat   5280
ttttatgttt tttgttgtta ttttaagtaa gaggaaaaat aggtttgttg tgaattatag   5340
ttttatggtt aaaatagaat gttagttaaa agtgtatgga tattaagttt ataaaatagg   5400
atatgggtgg tttttttgaa agaatataat ttaataataa aagtttttttg ggatatatgt   5460
ggattaaatg ttttattggt tttagttttt agttttttgaa tagaggtatt gttatgtttt   5520
tgattgtatt aggaaattag attttggaat aaatgttttg gtatttttagg gatgtgtttt   5580
tagttgaaat gtaatatttt tttatttttgt taattaaatt tataaattaaaa gtttaggata   5640
gtttagttga ttgttttttg taaatatgtg aaaaatatgt attattaaaa gtttaggata   5700
tgaatataag ataaaggtag atattttttgt tttaaattga tttttaggttt ttgagttgta   5760
ttgattgtga ttgggaatga ggtttttttat tttttagtgg tgtttatatg gatttatttt   5820
ttgattattt tttttaattat tttttttggtt atagtattaa tatgtaatat tgaggtgttt   5880
```

```
ttagagtgtt tatgtttagg gttttaggat atatgatatt taatggaggt tttgttgtta      5940
gatattagtt attattttgg atattaaatg attttaatta taatgttagg agtggttggt      6000
ttttggtttt gggatattat gtagttattg agagatttat gagtggtttt tgagattagt      6060
ataaaaaga aatagaaagt tataaaaatg ttaatgatgt tattatgtaa atatatgttt       6120
ttgtgttttg aaagattttt agtattgtag tgtttgagta taggagagtt ttttttatag      6180
ttagtattga aaataaatat tggatataaa taaatattga aaagaaagat tgttattttt      6240
tgttggtgat agtggtgttt tttgtaggtt aataatggtt atttatgttt tagattagtt      6300
ttagaaaaaa gtaagagtat ttagggaggg aggagagagg aataggggaa aggagaagga      6360
aaggaaaggg gatttgtaat tgtttattat tgatatagga agaataagaa ggttagttgt      6420
tttttttatag gttttgattg tttagagatt tataattaaa gttagtttaa gaagtttagt      6480
aaaggtagtt tttttaatta tttttttttt agtatttttt tttaaattta ttttggtgag      6540
ttttgttttt gggtttggtg agtatgggtg ggaaagtata ttgtgttatg ttgatttttt      6600
tttttttgttt ttggtaaaaa ttttatttgg gttttttagtt ttttggttt ttttttttat      6660
tttttttttgt tggatatata tgttttagtt gtgttttta gttttatttt tttagtggtt      6720
ggtaaagatt gatttggttt ttttgttgaa tgtgttttta gtttttatttt tttagtggtt      6780
tgttgggtat tgtaggtttg tttaaatatg agttattatt tttaaatatg gtgttttttt      6840
taattttttt gtgtttgtgt ttagtattat tgtttttta gttatttaga ttaaaatttt       6900
attgtttttt ttgagggttg attttttttt gttttttata tttaattaag aggtaatttt      6960
taagtttttt agttttataa tttttttttt tttattgtat t                          7001
```

<210> 24
<211> 11001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 24

```
ttaagttaga tgttttttaa ttatttgtgg ataggttagg tatatttttga gtttaatttt      60
attttatagg tttaatatttt tggagttaga aagttttttag gtaaaaagtt tgaagggggt     120
ttttttatgt tattagatgg attttttgtat tttttagaaga ttttttatat taggaaagat    180
taaagtatta aggtaatttt ttttggtttt ttgggataat tttaggtttt ggtatgagtg      240
gtttggaagt ttttgtttta gttataatgt ttatatattt ttggaattgt tttgtagggt      300
ttgttttta gtataatttt ttttttaagt tttattgtag ttatagttta ttagttttgt       360
ttagtgataa ttaagaaatt aagaattatg tatttatgtt tattttttta gagttatttt      420
tttttaggat aaagtttagg gttttgttat tgggtttttgt taggagttgt agttgtaggg      480
gttgtttatt atttaatagt ttttgtaagt atattgtgaa ggggaagtaa tgattagaga      540
tagggttagt tgtttagttt ttgtatgttt aggtgtatgt gtatatattt ttatataggg      600
tagggtgggg tgggaagttt attttggttg tgatgtttag tgtgtttaaa gagtgtaaat      660
ttgtggggggt tatttattat taagaatttt tgtagtaggg ttttaatgat tttatgtgtt      720
tttgtgtgtg atagtatatg taggtgtttg atagtatttt tgggtaggta aaaggaagtg      780
tggtgtttga tatttgtgtt ttgttttggg tattttgtaa ggagggtggt                  840
tttttgttgg agggtataga gatagggtgt agagaaatta aattttgata ataggtttta      900
gtgtaagaat tgttttttgtt ttaaagaaat agagaaatta aattttgata ataggtttta      960
ggtgagatgt tagtttatttt ggggttaggt tgggtatgta taaattattg tttgtgttta     1020
tttaagataa ttttagttgt gattttttga gtattaggta tatagttggg tatttgtttt     1080
ttttttatgtt tttttttttga gtagttaatt tattaagtttt atgaagaggt tgttgttgat    1140
ttgggtattg tatttttttga tttttttgttt aatttttagtt tgagaaaggt taggtgtttt   1200
ttattttata ggtttgtttt gtaagatggg ttagtatgga tatagggttt ttgaggaatt     1260
tagggttttt ttgaaaaatg gtttttgggg tagttttttgg aaattgattg ttttttggttt   1320
tttgtttttg atgtatatat atatagttgg tgtttatttt gaatttatta ttgttttttgg    1380
ttttgtatgt tttgggtgga taagggaaag atagaattat ttggttttttt tttgttgttt    1440
gtttagggtt ttagtattga atgtagtttt aaggatatta tagaagtagg ggtaattgaa     1500
ggtatatggt taggggttag gaatagttga gggattttga agagggattt ttatttaaag     1560
taaaattagg ttgggtgtgg tggtttatat ttgtaatttt agtattttgg gaggttaagg     1620
taggaggatt atttgatttt taggagtttg agattagttt gggtaatata gtaagatttt     1680
attttttatta aaaaaagaaa aaaaaaatt agttaggtgt ggtggtgtgt ttgtagtttt     1740
aattgtttag gaggttgagg tgggaggatt gtttgagttt gggagattgt agttatagta     1800
agttattatt gtgttattgt attttagttt ggggaattga gtgagatttt gttttaaaat     1860
ataaaaaata aaaataggtt gggtatgttg gtttatgttt gtaattttag tattttggga     1920
ggttgaggtg ggtggattat ttgaggttag gagtttgaga ttagtttgat taatatggag     1980
```

```
aaattttgtt tttattaaaa atataaaatt agttaggtgt ggtggtatat gtttgtaatt    2040
ttagttattt aggaggttga ggtaggagaa ttgtttaaat ttgggaggtg gaggttgtag    2100
tgaattgaga ttgtgttatt gtattttagt ttgggtaata agagtgaaat ttggttttaa    2160
aaaaaaaaaa aattagtaaa attatatttt aattgtatat tttgattata gtattttagt    2220
tgagttggag tgagggtttg ttttggagaa ggtagtttat ttttttttttt tgttttggta    2280
tggggttatg atttattgta gggtgagagg agtggagagt ggtgtatatt agtagtttag    2340
ttattagtgg atagagtagt atttggagtt agttttttat gttttatata tagtgagaaa    2400
aattattgtg atatgatgtt taattttgat ttaagttgta taaaaggtag ttttaggtta    2460
ggttttaatt tgttagaggt atataggtag ttttttggtg ggttttttgta tttgtttgtg    2520
ttgtttggag atttggttta aagatttttt ttttttttga gatgaagttt tattttgttg    2580
tttaggttgt agtgtagtgg ttggattttg gtttattgta agttttgttt tttgggttta    2640
agtgattttt ttgtttttaga tttttggagta gttgggatta taggtgtgtg ttataataat    2700
atttggttaa ttttttgtatt tttagtagag atgggatttt attatattgg ttaggttggt    2760
tttgaatttt tgatttttaag tgatttgttt gttttttattt tttaaagtgt tgggattata    2820
ggtgtgaatt attgtatttg atttagagat tttttaatttg attatttatt ttttatttta    2880
tttagggatt ggatttttgt tggaagggtg gagtgtggga tagggtagtt agggttttga    2940
attgattttt ttttttttaga ttttttttggt tttattgtat tagtttttatt ttttgttgat    3000
gttagatagg tttttagttag aatgtgagtg ttatagatat agttaagttt agtgttgatt    3060
aatattttgt tttagaagaa tttttataag gttttttgtta gaatgattttt gtgtttagtt    3120
taggagagtt agggtttttt ttgatttttgt tttggagttt ttttaagtat ttaaattatt    3180
tgatggggat aaatggagag gatagatgag ggagtaggggt ggagtgtttt agaattaag    3240
tttttatttta gaatttgttg ttatttttgta gttagtaagg atgtggggtt aagaattaag    3300
gttagggttt tataggaaaa aggtaaaggg ggaggggtgg gaatttaagt ttattttttt    3360
ttttaagtat ttaaaggttt tttggatgga gaagagtatt ggagtaaaaa ttttagtata    3420
aattttattg gggatagtgg gtaattttgt tgggttagta aaaataaatg gtgtgggttt    3480
tggaaaatga gggttggagg ttgtgaataa agtagtggat gtgtttgttt agtatattaa    3540
tgggaagaag tatttagatg ggaggagtat taggggtagg agaaatgtta gatagatttt    3600
agtgttaggg taagaaggaa gattattttg tttgtagaat agggagggta tagggatggt    3660
gttaatttgt ttttgtgatg gttttgagtt tttatttaat aatgagaaag tttgtttttt    3720
tttttttttt tggatgattt aggagttttg ggttgggatg tagtgatttt attttttagtt    3780
tttttttttt tggtgatgaa ttttttttatt tttatttaga aaatagattt ggattagagg    3840
tattgtatag ttttttttagg attttaaagg aggaagagtt ttttttttttg ttttttaaagt    3900
tgtttgttgg aagaggattt taatagttat tttagttgga tgtatagtag gattatggaa    3960

ttttttttttt gtattatagg gatttatttt ttatttttatt attgtttata aaaattgatg    4020
gtttttttttt tgagatagag ttttgtttttg ttttttaggt tggagtgtag tggtgtgatt    4080
ttggtttatt gtaattttttg ttttttgggt ttaagtaatt ttttgtttta gttttttaag    4140
tagttgggat tataggtgtt tgttattata attggttaat tttttgtatt tttagttgag    4200
atggggtttt attatttttgg ttaggttggt tttgaatttt tgattttatg atttatttat    4260
tttggttttt taaagtgttg ggattaaagg tgtgagttat tgtatttggt ttaaaattga    4320
tgtttttttt ttttttttta atatataatt tgggattttt tagtttttta ttttttttttt    4380
tttttttttt tttttttga gatagagttt tgttttttta tttaggttgg aatgtagtgg    4440
tttagttttg atttattgta attttttgttt tttgggtttta agtgatattt ttgtttttagt    4500
ttttttagta gttgggatta taggtatata ttattatggt tagataattt ttttgtattt    4560
ttagtataga tggggttttg ttatgttggt ttggtaggtt ttgaattttt ggttttaagt    4620
gatttgtttg ttttggtttt ttaaaatgtt gagattatag gtatgagtta ttaagtttag    4680
ttttttttttt tttttttgag atagagtttt attttgttat ttaggttgga gtgtagtggt    4740
atgatttttgg tttattgtaa tttttgtttt ttggttgaag tgtatgttta gtagagatag    4800
ttgggattat agttatatat tattatgttt ggttaattttt tgtaaatgat ttatttgttt    4860
ggtttttatta tgttggttag gttgattttg aatttttgat tgtaaatgat ttatttgttt    4920
tggtttttta aagtattggt attagaggtg tgagttattg tatttggttt tttttttttat    4980
ttttgagata gagtttttatt ttgttatttta ggttggagtg tagtggtatg atttttggttt    5040
attgtaattt ttgttttttta ggtttaagtg attttttttgt tttatttttt taagtagttg    5100
ggattatagg tgtgtatttt tgtggttagt tttttttttttt aattggttag tgttttgtgg    5160
tttttttatt tttttatagt ggaaaatggt ttaggattga ttgatatgaa gataagttta    5220
ggggtttata tttaatttaa tttttgtatt taagttttgg gttaagattt ggtgtgttg    5280
agtattattt attttgtaag gaattttgta aaattttatt tgaagtatta tttataattt    5340
tatttttttt atttaaataa ggattttttgt tttatttttg ttaggtatat tgagtttttat    5400
agttttttgtt tttttttttt ggtgtttagg tttggttttt tgagtttggt ggttatatta    5460
atggtatttg gtatatagtt ttttgataat ggggatattt aggaggtttt gagatatttt    5520
atagttttgg gttagtaatt tggattttttt tttttatttt tttaggtatt ttataattta    5580
gtttttttttt tttttgtggg taaagtgttt ttgaatgttt atggtttaaa ataagatttt    5640
ttttttattt attttttaaat ttttttttag atttattttta gaggaaggga atagaatttt    5700
```

```
ttatatttta gtagttggtg ataggttaga atagggaaga ggtgagggtt tagttggttt    5760
tatataggag tgtagatgga ggagtaggat tttttttttgt tttttaagtt tttttaaata    5820
tatttttaa tttttggtga ggattttttt tttttttatat tttttttttag tttttttaag    5880
gagggagtag gagtatttga atgtggaaat tgaggtgtta gtttaaattg tttggttggt    5940
tttagttata gttggataat gtttggttta ggtttattat aagttatata gttgttttt    6000
ttgtgtttaa tttgtttgtg atagaaatta agggggtttt ggtatttagt atttaggtgg    6060
tggaattggg gtttttatgta tggttttgtg ggtaggtttt tggttaggat ttgtggggag    6120
ttatgtagtt aggagggtgg ggttgtttat tgatttagga tgtggtaatg gattggggag    6180
ggtggagttt tagtgattgt tttttttttt gtttgttggt attttttggt ttttatttgg    6240
ttttggtgtg gtttgtgagt tagtgaggtt tgtgtggtga agtattgttt gagtttgag    6300
tttgagtttt tttggttgta gtagttattg tttgttgtgt tgttttaggt tattttgaaa    6360
gaaggtgttt ttgttttgtt tatagttgta tttgtttgtt ttttagtttt gtgtgtttgt    6420
agttgttaat tattgttttg gttgtgtgtg tgtgtgtatg tgtgttagtg tgtgtgtgtg    6480
tttgggttag agttgtgttg taattgttaa gattgaaatg tagattgttg ggatttagtt    6540
tttgttttat tggggtagga atgttggggt ggggatatgt atgtttttgtt tttaggaatg    6600
attttattgt tttggagttt tatttataga ttttatttat tatagggaat ggggggtgggt    6660
gttagtgttt gggtaagtgt ataagagtgg ttttttggttg gaggtgaggg tgggaaggtg    6720
tgggaagtgt gtgtgtgtgg agtttgggtt agtttgggtt tgggtttgtt tgtagtgggt    6780
ggagtatttg tggagttggt aatttaggtt ttttttttag tttttgtgta gaattagttt    6840
ttttgtgttg ttgggaaatt ggtaattaga atgttttttg tgtgtggtat ttaggtagtt    6900
tttgagaatg tttgtattgt ggtttgttta tttttgtttt ttttatatgt ttttggtttt    6960
gtgtttatta tgtttgtgtt ttgttttatt gtgggttttt agtttaggtt ttggggtttg    7020
taatagttta ggtagatgag tgtgtggtag tggtagtggt aggtgaattt gtaatttgta    7080
gagaggtttg gtggtgaggt ggaggagttt taggttgggg aaatgttttg gagattgaag    7140
ggaagtttta gggagagggt tgttgtttgt taggttttgt aggtttgatt tattttagtg    7200
ggttatttta tattgttatg tggatttttaa tgttggttat ttgggtgttt ggaaattggt    7260
tggaaggtta taggtagaga ggtttgttta atagttggat tttttattgtt tagtatagaa    7320
tttttttttt ttttattggt aattaaaaaa ataataataa aaaattgtgt tttgttttttg    7380
ttatttaggt tggagtgtaa tggtgtgatt ttggtttatt gtaatttttg tttttttgggt    7440
ttaagtgatt ttttttgtttt agttttttga gtatttagga ttataggtgt ttgttattat    7500
gtttagttaa ttttttgtatt tttagtagag atggggtttt attatgttag ttaggttggt    7560
ttttaaatttt tgatttttagg tgatttatttt gttttggttt tttaaagtgt tgggattata    7620
ggtttgagtt attgtatttg gttttttttatt gggaatgtat atggaatata tttgtttatt    7680
tatttgaagg aaaaattaaa tatttttaat ttatgtttgt tttgtggttg ttatttgttt    7740
ttattttttt tagattaaga tattggtttt tatatatttt aattttttgt ttttatttttt    7800
ttttttattt attttagtta ggtttgtttt tttttttagg aaattgtttg ggatagggtt    7860
tttagtgatt tgtgtattat taaatggaat ttagtgtttt attttttatt tttattttttt    7920
tagtattatt tgaagtttgt tttttttgat ttttaggggt tatattttttt tagtttttttt    7980
tttatttttt gtagtttttg tttagttttt ttgtagattt tgatttaatt tttatatttt    8040
atgatgaagt ttgggtttag ttttgattat tggtttggtt tgtttatatt tatttgtttt    8100
agatttatgg ttgaaatatt gatttaaatt tttagattag atttttttgtg ttagtattttt ..    8160
tattaggatg tttaaaagat gttttaagtg aatatggtta aaatttaatt tttttttttt    8220
tagttttatt gttatatttg tttagttttt ttttttgtagt aaaaatggtt attaggtttt    8280
tagttattgg agataaaagt ttaaatttat ttttgatttt ttttttttgtttt ttattttttga    8340
taaatatgtt taaattattt tgttttttat ttttatggtt attttatttt tttttgagaa    8400
tgttgtaatg ttttagtttt gtttttttttt ttttttttttt ttttttttgag atagagtttt    8460
attttgttgt taaggttgga gggtagtggt atgattttgg tttattgtaa tttttgtttt  .  8520
ttgtgtttaa gtaattttttt tatttttagtt ttttgagtag ttgggattat aggtatttgt ·   8580
tattatgttt ggtttatttt tttttatttt ttagtagata tgaggtttta ttatgttggt ·    8640
taggttggtt ttgaatttttt gatttttaggt gattttttttt atttttgtttt ttgaagtgtt    8700
gggattatag gtatgagtta ttgtgtttgg tttttttttg tattggatga gagggttttt    8760
attttgtgtt ttttttagtt gaattgtgga tgtttttttg ttaaataggt tagagtagtg  -    8820
atgtatatat agatttggga tattatttat ttataagttt ttaaataggt tattgatgtt    8880
atgtttaatt tagatttttat gttataataa tttggggagt ttttaaaatt tattgatgtt    8940
tagggtttat ttttagtagt tgatttaata ggtttgtggt gggatttagg ttagtgggga    9000
ggattgtaaa agtatttttg gtgatttttag ttggtgttta tttaggggag agtaattttt    9060
gtttgttggt gattttttagg ggtgtagaag gattgttggg tgtgtggttg tgtgtatatt    9120
ttagtatttg atttattggg ttagaaaagg gtgtttgtta aataaagatt aataaaatt    9180
tttgtttgta gggggtttat taaaggtttt aaattttttt aggtttttttt ttataggtgg    9240
taattttttt ttattttaaa ggttttggag ggggttatga gtgtttgaga agaggtaagt    9300
ttgggaagat ggattttgag gatagtaggt ataaattttt ttttaagaag ggttaaggta  ··  9360
tttttaaagat aagaaattta aaattagtgt attttttatat ataagtagtt attttttgttt    9420
atttgtggtt tagatatgag tggagtgtga taagggataa attatttttg tgtattttttt    9480
```

```
agtgatgggg tgaaagtaat ggatttagtt tttgggagtt gtttttgttg attttttttg   9540
ttgtgatttg atttgtggtg attgtgttgt tttttggttg ttttttttgt ttttgtaggt   9600
gtgtggggtt attatttatg tgtgtattgt aggttttgt gtatgatgtt ttagatgaag    9660
ttgttataga ggttgtatta tgtgtgtgtg gtgggttttg tgggttggaa gtggtggtta    9720
tggttaggga ttagttgttg tgtgggggttg tatgtggtgt tttgtgtgat gtgtagtgtg   9780
ttggtatgtt ttagtttgggt gtggtttttt ttagtgtgtt tagtgggtgt tagtttttgt    9840
agtttaatga gtttaggttt ttttgatatg gtttggttgg gtttgtgttt tgttggtttt    9900
gggtgttagt aagtgtgggt tgggtgggt tataggtgg gttttgattt tagtgttttt      9960
tttaggattt agattgggtg gtgggaagga gttgaggaga gttgtgtaat ggaaatttgg   10020
gtgtagggat tgtggggttt gaaggtgggg ttgggtgtgt ttttgtagag ttttttttgt    10080
tttgtttttt tttttttttt ttgttttttt tttatatttt attttggatg gttataatga    10140
tggtgattgt aaagtattat gtggagatat ttgtgttttt ggaggttagt tttattgtgt     10200
tagaggaaga gggttttat atttggtttt ggttttttg gtttggtttg ttgaagtaat       10260
atatttggtt tatttattgg gtggggtagg aagttttgag tttttatttg gggtgaggag    10320
gagggagatt ggttagtagt tttattgttt gttttgtttt ttattgtgga gattgggtt      10380
ttggtagagg ttggattgtg attttgaggt ttaggggtgt attttgggtg gatttttttg    10440
gtatgggtgg ttggtttttta gtaattgtag ttttttatttg gttttgttat tttgggttgt    10500
taggatataa gttttttta gttttttta gtgtttgatt tggtattttt tgtaggtagg       10560
tgggtattga ggatggtaat gtatgtgggg gatgtgggag taggtgtttag aggtttaagg   10620
tttttaggata ttttttattttg tagtaatatt atttattttg gtattgtgag tagtgtttag  10680
aagttttgt attgtagtaa gtatagtggg gttgttttgg agttattgtt tttagtatat      10740
ttagtttgta ggttttagtt tatttggggg aaagttagga aggtttgatt ggttttggaa     10800
ggtggggta ttttatttat atttatgttt tttgtattttt ttttattttt tttgttatttt    10860
ttataggttt tattttgtg tttgtagttg taggttttgt tttgaggggt tgaatatatg      10920
ttggagttgg tgtttggtaa ttgtttgtta tttgttttg tttttttgtt ttagttgttt      10980
ttagattttt gggatttagg a                                               11001
```

<210> 25
<211> 11001
<212> DNA
<213> Artificial Sequence

<220>
<223> chemically treated genomic DNA (Homo sapiens)

<400> 25

```
ttttagattt tagaaatttg ggagtggttg gagtgagaaa atagaggtaa gtggtaggta     60
attgttaagt attagtttta gtatgtgttt agtttttag agtaggattt gtggttgtag     120
gtgtgaaggt aaggtttgtg gaaatggtag ggagggtgga ggggatgtag gaggtatgga    180
tgtgggtggg gtgtttttat tttttagggt tagttagatt ttttttgattt tttttttaggt  .240
gggttgagat ttataggttg gatgtgttag aggtagtggt tttagagtgg ttttgttgtg    300
tttattgtag tgtagaggtt tttaagtgtt gtttatgatg ttagaatgag tggtattgtt    360
gtaggtgagg gtattttaga attttggatt tttaagtttt atttttatat tttttatatg    420
tattgttatt tttaatattt atttgtttgt agggagtgtt aagttaagta ttgggaaaag    480
tatggaaaga tttgtgtttt ggtagtttag ggtgatagag ttaaatgagg gttgtagttg    540
ttgagggttg attatttatg ttaagggaat ttatttagaa tgtatttttg aattttaaga   600
ttatggttta gttttttgttg gagttttagt ttttgtagtg gagagtagag tgggtggtaa   660
agttgttgat tgattttttt ttttttttatt ttaagtgaag gtttgagatt ttttgttta    720
tttagtgggt aggttaagtg tgttgtttta gtaaattgga ttaggaggt tagggttgga     780
tgtgggatt ttttttttttt agtatagtaa agttggtttt tagaaatatg ggtattttttg   840
tgtggtgttt tgtggttgtt gttgttgtgg ttgtttgggg tggggtgtga ggagggatg     900
aaggagggaa ggaagggtaa ggtgggggg gtttgtgag agtgtgttta gttttgtttt    960
tgggttttat agttttgta tttaggtttt tattgtgtgg ttttttttag tttttttttg    1020
ttgtttagtt tggatttttgg gggaggtgtt gaagttgggg tttgttttgt ggttttgttt    1080
ggtttgtgtt tgttagtgtt taaagttagt gaagtatggg tttaattggg ttatgttggg    1140
ggagtttgag tttattgagt tgtgggagtt ggtatttgtt gggtgtgttg ggaagggttg    1200
tatttggttg gagtgtgtta atgtgttgtg tattgtgtgg ggtattgtgt gtaattttat    1260
atggtagttg gttttttggtt gtggttattg tttttagttt gtgggtttg ttatgtatat    1320
gtggtgtgat ttttgtggtg attttatttg gggtgttgtg tgtaaaggtt gtagtgtgt    1380
gtgtgagtag tggttttgtg tgtttatgag agtggaaggg gtagttaagg ggtagtgtag    1440
ttgttgtggg ttaagttgtg gtagaggggg ttggtgggga tagttttga ggattaggtt     1500
tgttattttt gttttattgt tgaagagtgt gtgaaaatgg tttattttt gttgtatttt    1560
```

```
atttgtattt gggttataga tgagtagagg tggttgttta tatgtaaaaa tatgttgatt    1620
ttaagttttt tattttaaa atgttttggt ttttttgag aaagggtttg tgtttattgt    1680
ttttggagtt tatttttta ggtttgtttt tttttaaata tttatgattt tttttagaat    1740
ttttaggrtg aagggaaatt attatttatg ggagggagtt tggaaaaatt tagaattttt    1800
ggtgggtttt ttgtaagtag gagtttttgtt gagttttat ttagtaaata ttttttttttg    1860
atttagtgaa ttagatgtta aaatatgtat gtagttatat atttagtagt tttttttgtat    1920
ttttgggaat tgttagtaag taaaggttgt ttttttttgg gtagatatta gttggaatta    1980
ttaggggtgt ttttatagtt ttttttgtta gtttggattt tattgtagat ttgttgaatt    2040
aattgttggg agtggatttt aggtattagt aaattttaaa aatttttttaa attattgtaa    2100
tatggagttt gggttgagta ttattgtttt ggtttatttta ggaatttgtg gatggatagt    2160
gttttaggtt tgtgtgtgta tggagatttt tttatttggt ataagaggat attataaatt    2220
tagttggggg gagtataaag ttgtgataga atgtaaagaa tgaataaggg gttgagtgtg    2280
gtggtttatg tttgtaattt tagtattttg gaaggtggag gtgggtggat tatttgaggt    2340
taggagttta agattagttt ggttaatatg gtgaaatttt atgtttatta aaaaataaaa    2400
aaaaatgagt taggtgtagt ggtgggtgtt tgtaatttta gttatttggg aggttgaggt    2460
gggagaattg tttgaatata ggaggtggag gttgtagtga gttgagattg tgttattgtt    2520
ttttagtttt ggtgatagag tgagatttttg ttttaaaaaa aaaaaaaaaa aaaaaagaa    2580
taaggttggg atattgtagt gtttttaaag agaaataaag tagttatgga gataagaagt    2640
aggatgattt gggtatgttt attagaggta gagataaggg agaaattaaa gataagtttg    2700
ggttttttgtt tttagtaatt gggagtttag tggttatttt tgttgtaaag aggaagttgg    2760
gtaagtgtag tagtgaggtt gaagaaaagg gaattaaatt ttggttatgt ttatttgaaa    2820
tgtttttttag atatttttagt gaaggtattg gtatggagga tttagtttga gggtttaggt    2880
tagtgtttta gttgtggatt tggggtagat gaatgtagat agattaggtt agtgattagg    2940
attgagttta gatttttattg tgagatatgg aagttgagtt agaatttgta aaggagttga    3000
gtaggagttg taggggrtag gaggaaaatt gggagagtgt agttttttggg agttaaaggg    3060
agtaagtttt aaatgatgtt gagggggtga gaatggagaa tggaatattg gattttatttt    3120
ggtagtatat agattgttga ggattttgtt ttgggtagtt ttttggagga agaggtaagt    3180
ttggttggag tgggtagagg ggagagtgaa ggtgaaggat tagagtgtat agagattagt    3240
gttttggttt gagggggagta gagataggtg ataattatag ggtagatgta ggttaaaggt    3300
gtttagtttt tttttttaagt aaatgggtag atgtatttta tatatgtttt tagtgaaggg    3360
ttgggtgtgg tggtttaagt ttgtagtttt agtattttgg aaggttgagg tgggtggatt    3420
atttgagatt aggagtttga gattagtttg gttaatatgg tgaaattttg tttttattaa    3480
aaatataaaa attagttggg tatggtggtg ggtgtttgta attttaggta tttaggaggt    3540
tgaggtagaa gaattgttttg aatttaggag gtggaggttg tggtgagttg aaattgtgtt    3600
attgtatttt agtttggggtg ataaaagtaa gatgtagttt tttgttgttg tttttttaat    3660
tgttaatgag gaaaggggaa gttttgtgtt aggtgataga gatttaattg ttgagtaggt    3720
ttttttgttt gtggttttttt ggttggtttt tagatgttta ggtggttaat attagagttt    3780
gtgtagtagt gtgaggtaat ttattgagat aggttgggtt tgtggagttt ggtgagtagt    3840
ggtttttttt ttggggtttt tttttaattt ttgggatatt tttttgattt ggagtttttt    3900
tgttttattg ttaggttttt ttgtagattg taagtttatt tgttattatt gttgttgtgt    3960
gtttgtttgt ttggattgtt gtgggttttg ggatttgggt tgggaatttg tggtggagtg    4020
ggatatgaat gtggtgagtg tggggttgag ggtgtatggg aagggtgagg atgggtaggt    4080
tatagtgtag gtatttttga gggttgtttg ggtgttgtgt gtaaggagtg ttttaattgt    4140
tgatttttttg gtggtataga gaggttaatt ttgtgtgggg gttgggaggg gagtttggat    4200
tgttggtttt gtaagtattt tatttgttgt aagtggattt gggtttaggt tgatttaggt    4260
tttgtgtatg tgtattttttt gtattttttt gtttttgttt ttggttagag gttatttttg    4320
tgtgtttgtt tggatgttgg tatttgtttt tgtttttttgt ggtaggtggg gtttgtgagt    4380
ggagttttgg agtgatgagg ttatttttgg gggtgaagtg tgtgtgtttt tgttttggtg    4440
tttttgtttt aatgagataa gagttagatt ttggtgattt atgttttagt tttaatggtt    4500
gtggtgtggt tttggtttgg gtgtatgtgt atattgatat gtgtatatgt atgtatgtga    4560
ttggggtggt ggttggtggt tatggatgtg taggattggg ggatgggtgg gtatggttat    4620
gggtgaggtg gaggtgtttt tttttgaaat gatttggagt agtatgatga gtagtggtta    4680
ttgtagttaa gaggatttgg atttggagtt tgagtagtat tttattgtgt gaatttttgtt    4740
agtttgtagg ttgtgttggg attaggtggg agttaggggg tgttggtggg tgggaggggg    4800
agtggttgtt ggagtttttgt ttttttttggt ttgttgttgt gtttttgggtt ggtgggtagt    4860
tttatttttt tggttatgtg gttttttttgtg ggttttggtt ggggatttgt ttgtggaatt    4920
gtgtgtaaga tttttgattt attgtttaga tgttgggtgt tggggttttt ttggtttttg    4980
ttatagatag gttgaatatg gaaaaagtag ttgtatggtt tgtggtagat ttgagttggg    5040
tattatttag ttatgattaa agttgattga gtagtttgga ttagtatttt gattttttgtg    5100
tttgaatgtt tttgtttttt ttttggggag attagggggag gatgtggaga gggaagagtt    5160
tttgttagga attgagaagt atgttttagga aaatttgaga ggtagagaga gatttttgttt    5220
ttttatttgt attttttgtat ggagttagtt gagtttttat ttttttttttg tttttggtttg    5280
ttattagttg ttggaatgtg gaagattttg tttttttttt ttagggtgga tttggagaaa    5340
```

```
gatttgggaa tagataggaa agaagttttg ttttggatta taagtattta ggagtatttt    5400
atttatagga aggggaaag ttagattata aaatgtttaa agaggtggaa aaagagattt     5460
aggttattaa tttaggattg taaggtgttt tggaattttt taggtatttt tattattgga    5520
gaattgtgtg ttagatgtta ttggtgtgat tattaggttt agagaattag gtttaggtat    5580
taggaaaaag aaataggggat tgtgaagttt agtatgtttg gtagaaatgg ggtggaaatt    5640
tttatttaag taaagaaagt ggagttgtga gtgatgtttt agataaaatt ttataaaatt    5700
ttttataaaa tgggtggtgt ttagtatgtt aaaattttag tttagagttt gggtgtaagg    5760
gttgagttga gtgtagattt ttgggtttgt ttttatgtta gttagttttg agttattttt    5820
tattgtggaa aggtgggaaa attataagat attaattaat tgaaaaggag ggttagttat    5880
ggaggtgtat atttgtaatt ttagttattt gggagggtga ggtagaagga ttatttgaat    5940
ttgggaggta gaggttgtag tgagttaaga ttgtgttatt gtattttagt ttgagtgata    6000
gagtgagatt ttgttttaaa aatagaaaag gaagttaagt atggtggttt atatttttaa    6060
tgttaatgtt ttgggaggtt aaggtaggtg gattatttgt aattaggaat ttgaggttag    6120
tttggttaat atggtgaaat tttattttta ttaaatatat aaaaattagt tgggtatggt    6180
ggtgtgtgat tgtagtttta gttatttggg agattgaatt attttaattg ggaggtaaag    6240
gttgtagtga gttaagattg tgttattgta ttttaatttg ggtgataggg tgaggttttg    6300
ttttaaaaaa aagaagaag gttgggtttg gtgatttatg tttgtaattt tagtattttg     6360
ggaggttaag gtaggtagat tatttgaggt taagagtttg agatttgtta ggttaatata    6420
gtaaaatttt gtttgtattg aaaatatäaa aaaattattt ggttatggtg gtgtgtgttt    6480
gtaattttag ttattgggga ggttgaggta ggagtattat ttgaatttag aagatagagg    6540
ttgtagtgag ttgagattgg gttattgtat tttagtttgg atgagagagt aagattttgt    6600
tttaaaaaaa aaaaaaaaa aaaagaaaga ataggaggtt gagaagtttt aagttatatg     6660
ttaaaaaaaa agaaaaaaat attagtttta ggttaggtgt agtggtttat atttttaatt    6720
ttagtatttt ggaaagttga ggtgggtgga ttatgaggtt aggagtttaa gattagtttg    6780
gttaaaatgg tgaaattttg ttttgattaa aaatataaaa aattagttag ttgtggtggt    6840
aggtatttgt aattttagtt atttgggagg ttgaagtaga gaattgtttg aatttaggag    6900
gtagagattg taatgagtta agattgtatt attgtatttt agtttggaaa atagagtgag    6960
atttttgtttt aaaaaaaaaa ttattagttt ttatggatag tggtagagtg gagggtgggt    7020
ttttatggtg tagaagggaa attttatggt tttgttgtgt atttgattgg gatggttgtt    7080
gaaattttt tttagtaggt agttttggaa atagaaaaag aaatttttt ttttttagaa      7140
ttttggaagg gttgtgtagt gtttttaatt taagtttgtt ttttgagtga agataggggag   7200
gtttattatt agaagggaag gggttggaaa tgaggttatt gtattttagt ttagggtttt    7260
tgggttattt aggaagggaa gaaggagtaa gttttttat tgttaggtag gagtttagag     7320
ttattataag aataagttag tattatttt gtgtttttt tgttttgtaa ataaaatgat      7380
tttttttttt gttttggtat tagagtttgt ttggtatttt ttttgttttt agtattttt     7440
ttatttgggt attttttttt gttggtgtat tgaataaata tatttattgt tttatttata    7500
gttttagtt tttattttt agggtttata ttatttgttt ttattaattt gataaggttg      7560
tttattgttt ttagtaaggt ttgtattggg gtttttattt tagtgttttt tttttatttag   7620
gagatttttg gatatttggg gaagaaaatg agtttaaatt tttatttttt tttttttatt    7680
ttttttttgt aaggtttggg ttttagtttt tagttttata tttttgttgg ttgtagaata    7740
gtagtgggtt ttgggtaagg agtattttgt taaaatgttt tattttgttt ttttatttgt    7800
ttttttatt tgtttttatt agatggttta agtgtttaag gggatttttag ggtggagtta   7860
gggagaattt tggtttttttt gggttaggta taagattatt ttataggaaa ttttgtggga  7920
attttttgg gataaagtat tggttagtgt tgagtttagt tgtgtttgtg atatttgtat     7980
tttaattagg gtttatttga tgttaatagg aagtaaggtt gatgtagtgg ggttaaggga    8040
gtttgggaga agaaagttgg tttagagttt tggttgtttt gttttatatt ttatttttt    8100
ggtaagaatt tagtttttag atgaggtggg gagtgagtgg ttgagttaaa aattttttgg    8160
ttgggtatga tggtttatgt ttgtaatttt agtattttgÿ gaggtgaagg taggtggatt    8220
atttgaggtt aggagtttaa gattaatttg gttaatgtgg tgaaatttta ttttattaa     8280
aaatataaaa attagttggg tgttgttgtg gtatgtgttt gtagttttag ttatttggga    8340
gtttgaggta ggagaattgt ttgaatttag gaggtagaat ttgtagtgag ttaagattta    8400
gttattgtat tatagtttgg gtgatagagt gaggttttgt tttaaaaaaa aaaaaaattt    8460
ttgggttaaa tttttagata gtataggtag gtgtagaaat ttattaggaa gttgtttgtg    8520
tatttttggt agattggagt ttggttttaaa gttgtttttt atgtagtttg ggttaaggtt   8580
aaatattatg ttatagtgat ttttttttatt atgtgtgaga tatggagaat tggtttttaag  8640
tattattttg tttattggtg gttggattat tgatgtgtat tatttttat tttttttatt     8700
ttgtagtggg ttatggtttt gtgttggggt agaggagaaa aatgggttgt tttttttagg    8760
ataaatttt attttaattt aattagggtg ttgtgattag aatgtgtaat tgaggtgtga    8820
ttttattgat tttttttttt tttgagattg agttttgttt ttgttgttta ggttggagtg   8880
tgatggtatg attttagttt attgtaattt ttatttttg agtttgagta attttttttgt    8940
tttagttttt taagtagttg ggattatagg tatgtgttat tatgtttggt taattttgta    9000
tttttagtag agatgggggt tttttatgtt ggttaggttg gttttaaatt tttgatttta    9060
ggtgatttat ttgttttggt tttttaaagt gttagaatta taggtgtgag ttaatgtgtt    9120
```

```
tagtttgttt ttgttttttg tgttttgaag tagggtttta tttagttttt taggttggag      9180
tgtagtgata tgataatagt ttattgtagt tgtaattttt tgggtttaaa tgatttttt       9240
attttagttt tttgaatagt tgggattata ggtatattat tatatttggt taatttttt       9300
tttttttttt ttagtagaga tgaggttttg ttatgttgtt taagttggtt ttaaattttt      9360
gaggattaag tgattttttt attttagttt tttaaaatgt tgggattgta gatgtgagtt      9420
attatattta gtttgatttt attttaaatg agagtttttt tttagagttt tttagttgtt      9480
tttggtttttt ggttatgtgt ttttagttgt ttttgttttt gtggtatttt taaggttata     9540
tttagtgttg aggtttttagg taggtagtag agagaagtta aatgattttg tttttttttt     9600
atttatttag agtatgtaaa attaggagta gtggtgggtt tagggtgggt attagttatg      9660
tatatgtata ttagggatag ggggttaaag gtagttagtt tttaaagatt gttttagagg      9720
ttatttttta gagaagtttt gggtttttta agggtttttgt gtttatgttg gtttattttg      9780
taggatgagt ttgtggagtg ggagatattt gatttttttt aagttgagat tgagtagaag      9840
attaaggagt ataatgttta gattaatagt aattttttta tgagtttggt gagttgattg      9900
tttaggaagg gggtgtgggg aggagtaggt atttagttat gtgtttgata tttagagggt      9960
tataattgag gttatttttgg gtgggtgtaa gtagtaattt gtgtatattt agtttagttt     10020
taagtagatt gatattttat ttggaatttta ttattaaggt ttggtttttt tattttttta     10080
gaataaggat ggtttttata taggttttat taaggtttag ttgaagttgg tgtgtttttgt     10140
ttttgtgttt tttagtaaga agttattttt tttgtaggat gtttggtggg gtttaggatg      10200
gggtataagt gttaggtgtt gtattttttt ttatttgttt aaggatgttg ttaagtattt      10260
gtatgtgttg ttatgtataa gggtatgtga agttattgag gttttgttgt gaaagttttt      10320
ggtggtggat gatttttgta agtttgtatt ttttgagtgt gttgagtgtt atggttaagg      10380
tgggtttttt attttatttt gttttatgtg agggtatata tgtatgtatt tgagtatgta      10440
ggggttgagt agttggtttt gtttttgatt attatttttt ttttatagtg tatttgtgga      10500
agttgttgga tgatgagtag tttttgtggt tgtggttttt ggtagggttt agtgataagg      10560
ttttgagttt tgttttgaag gaaaatgatt ttggggaggt gaatgtgagt atatagtttt      10620
tagttttttg gttgttatta gataggattg atgggttgta gttatagtaa ggtttggagg      10680
aggaattgtg ttggaagata agttttgtaa aatagtttta ggagtgtata ggtattgtaa      10740
ttaaagtaaa ggtttttaga ttatttatgt taaagtttag ggttgtttta agaagttagg      10800
aagaattgtt ttggtgtttt gatttttttt ggtgtggaaa attttttgga gatgtaggag      10860
tttatttaat gatatgagga ggtttttttt agatttttta tttggaagtt ttttggtttt      10920
aaggtattag gtttgtggag tgaaattaga tttagaatat gtttgatttg tttataggta      10980
                                                                        11001
attggggaat atttgatttg g
```

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP2004/011577

International filing date: 14 October 2004 (14.10.2004)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 103 48 407.8
Filing date: 17 October 2003 (17.10.2003)

Date of receipt at the International Bureau: 19 January 2005 (19.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record.

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☑ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☑ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☑ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.